



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 180386

TO: Jeanine Goldberg
Art Unit: 1634
Location: REM-2D15&2C70
Serial Number: 10/699941

Friday, March 10, 2006

From: Beverly Shears
Location: Biotech-Chem Library
REM 1A54
Phone: 571-272-2528
beverly.shears@uspto.gov

Search Notes

Your queries have completed processing. You may access an electronic version via eDAN (SCORE) and /or <http://es/ScoreAccessWeb>.

Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

Published Applications Database - November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).



THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

180386

From: Goldberg, Jeanine
Sent: Thursday, February 23, 2006 12:54 PM
To: STIC-Biotech/ChemLib
Subject: 10/699941- ataxia-

Hello-
Please search SEQ ID NO: 3.

THANKS
Jeanine

Jeanine Anne Goldberg
1634
571-272-0743
REM 2D15
Mailbox: 2C70

CRFES

RECEIVED
FEB 23 2006
STIC

3-2918NA

3/07
STB

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

Date completed: _____
Searcher: Beverly e 2528
Terminal time: _____
Elapsed time: _____
CPU time: _____
Total time: _____
Number of Searches: _____
Number of Databases: _____

Search Site

_____ STIC
_____ CM-1
_____ Pre-S

Type of Search

_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors

_____ IG
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
_____ Other CGN

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 03:15:24 ; Search time 14445 Seconds
(without alignments)

11482.810 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gcgcgcctctgcagccct.....tctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5881141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_in:*

3: gb_env:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pr:*

9: gb_ro:*

10: gb_sca:*

11: gb_sy:*

12: gb_un:*

13: gb_vl:*

14: gb_hlg:*

15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2918	100.0	2918	8	BC026217 Homo sapi
2	2896.6	99.3	4976	6	BD171164 Novel gen
3	2896.6	99.3	4976	6	BD183464 Novel gen
4	2896.6	99.3	4976	8	AB058775 Homo sapi
5	2887.6	99.0	3076	6	AR560886 Sequence
6	2870.2	98.4	2894	6	BD158743 Primer fo
7	2870.2	98.4	2894	6	AX881045 Sequence
8	2870.2	98.4	2894	6	AK027989
9	2598.2	89.0	2759	6	AX747472
10	2598.2	89.0	2759	8	AK092309
11	2463	84.4	2787	8	AB052149 Macaca fa
12	2357.6	80.8	2786	6	CQ842483 Sequence
13	2357.6	80.8	2786	8	AK125457 Homo sapi
14	1815.8	62.2	2276	8	BC008736
15	1438.2	49.3	134308	8	AC011488
16	1438.2	49.3	210617	14	AC034201 Homo sapi
17	1325.8	45.4	2625	6	AX747461 Sequence
18	1325.8	45.4	2625	8	AK092298 Homo sapi

19	1114.4	38.2	1116	8	AY220297
20	1113	38.1	1113	6	AR560887 Sequence
21	923.2	31.6	997	6	CQ721138 Sequence
22	857.8	29.4	1629	9	AY349150 Mus muscu
23	857.8	29.4	3677	9	BC048903 Mus muscu
24	745.2	25.5	790	6	BD150580 Primer fo
25	745.2	25.5	790	6	AX870518 Sequence
26	579	19.8	579	6	CQ832574 Sequence
27	522.2	17.9	1821	8	AB046623 Macaca fa
28	470	16.1	560	6	BD155501 Primer fo
29	470	16.1	560	6	AX875439 Sequence
30	346.4	11.9	4507	9	BC094224 Mus muscu
31	341	11.7	4164	9	BC095978 Mus muscu
32	332	11.4	2242	9	BC091368 Rattus no
33	330.4	11.3	5656	8	AB002365 Human mRN
34	327.6	11.2	11690	8	AB050197 Homo sapi
35	326.6	11.2	4284	9	AK172939 Mus muscu
36	318.2	10.9	5654	6	AX330135 Sequence
37	318.2	10.9	5654	6	AX336310 Sequence
38	318.2	10.9	5654	6	AX336690 Sequence
39	315.8	10.8	2347	6	BD083730 Nucleic a
40	315.8	10.8	2347	6	BD097377 Nucleic a
41	311.4	10.7	2491	8	AY439213 Homo sapi
42	311	10.7	1804	5	AY423003 Danio rer
43	309.8	10.6	322	6	BD118568 EST and e
44	309.8	10.6	322	6	AR423015 Sequence
45	309.8	10.6	322	6	AX983709 Sequence

ALIGNMENTS

RESULT 1

BC026217

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BC026217 Homo sapiens ataxia, cerebellar, Cayman type (caytaxin), mRNA (CDNA clone MGC:21400 IMAGE:4153341), complete cds.

BC026217 Homo sapiens (human)

MGC.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 2918)

Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Heiton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2918)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 2918)

Straussberg, R.

Direct Submission

Submitted (02-APR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Goldberg, J.
10/6/99 941
Seg. ID 3

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mhc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 cDNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdc@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 20 Row: b Column: 2
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

```

1..2918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:21400 IMAGE:4153341"
/tissue types="Brain, anaplastic oligodendroglioma with 1p/19q loss"
/clone_lib="NCI CGAP_Brn67"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..2918
/gene="ATCAY"
/notes="synonyms: KIAA1872, BNIP-H, CLAC"
/db_xref="GeneID:85300"
/db_xref="MIM:608179"
373..1488
/gene="ATCAY"
/codon_start=1
/product="caytaxin"
/protein_id="AAH26217.1"
/db_xref="GI:20070730"
/db_xref="GeneID:85300"
/db_xref="MIM:608179"

```

gene

CDS

```

1..2918
/gene="ATCAY"
/codon_start=1
/product="caytaxin"
/protein_id="AAH26217.1"
/db_xref="GI:20070730"
/db_xref="GeneID:85300"
/db_xref="MIM:608179"
translation="MGTTRATLMEVNDVKEDLDLPRPLPEETGVLLGSPVEDT
SSPNTLNFNGARKKTLVAPENISLQSGSLSDPFLTPDDLDINVDIETPD
ETDSEFLNGNELEWDDTPVATAKNMFGDSADLFGDTTDDGSAAGRLMKRTVIG
EQHRIDLHMRPYMKVTHGGYEGELNAIVFAACFLPDSLSLPDYHIMENLFLYV
ISLELLVAEDYIMVYINGATPRRRMGIGWLKCYQMIDRLRLKSLIIVHPSWF
IRTVLAISRPFISVKINKIQYVHSLDLLEQLIPMEHVQIPDCVLQYBEERLKARRS
ARQPEFVLPSEKEPVAPEVNSALVSEDQETSMS"

```

ORIGIN

Query Match 100.0%; Score 2918; DB 8; Length 2918;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GCGAGCCTCTGCAGCCCTGAGCTGGAGAGACGAGCTCTCGAGGCGAGGCGCGCA	60
DB	1	GCGAGCCTCTGCAGCCCTGAGCTGGAGAGACGAGCTCTCGAGGCGAGGCGCGCA	60
QY	61	GCGGCGGCGGATGAGAGGCGGCGCAGCCGCGCGCTGGGAGGCCACCGCTAAC	120
DB	61	GCGGCGGCGGATGAGAGGCGGCGCAGCCGCGCGCTGGGAGGCCACCGCTAAC	120
QY	121	CTTGCAACCCACCCCTGCGCAAAAGAGCTGGGCGGCGCTGGCGACGTCGCCCTGGG	180
DB	121	CTTGCAACCCACCCCTGCGCAAAAGAGCTGGGCGGCGCTGGCGACGTCGCCCTGGG	180
QY	181	TGACCTTCTTCGGATCGAGATCGCCCTCGGCGAGCATCTCTTCTCTCTAGCTTGAA	240
DB	181	TGACCTTCTTCGGATCGAGATCGCCCTCGGCGAGCATCTCTTCTCTCTAGCTTGAA	240

QY	241	GGCCCGGGGAGCGTGCAGCGATGCCAGCTGCCACCGGGCAGGGCTCGCTTGTGTGCCA	300
DB	241	GGCCCGGGGAGCGTGCAGCGATGCCAGCTGCCACCGGGCAGGGCTCGCTTGTGTGCCA	300
QY	301	GTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCACTCCCTGCTTCAAGCAGTGCCTC	360
DB	301	GTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCACTCCCTGCTTCAAGCAGTGCCTC	360
QY	361	TTCCAGCTCCCATCGGGGAGCCCAAGCCAGCTCCGGATCGGAAGAGCTGCAAGTGAAG	420
DB	361	TTCCAGCTCCCATCGGGGAGCCCAAGCCAGCTCCGGATCGGAAGAGCTGCAAGTGAAG	420
QY	421	GAGGAATGGCAGGACGAAGATCTTCCAGGCGCACTCCAGAGAGAGACGGGGGTGGAATG	480
DB	421	GAGGAATGGCAGGACGAAGATCTTCCAGGCGCACTCCAGAGAGAGACGGGGGTGGAATG	480
QY	481	CTTGGAGCCCGGTGGAAAGACACATCTCTCTCCCAACACGCTAAATTTCAACGAGCG	540
DB	481	CTTGGAGCCCGGTGGAAAGACACATCTCTCTCCCAACACGCTAAATTTCAACGAGCG	540
QY	541	CATCTGTAAGAGGACGCTGGTGGCCCGCAGAGATCAACATTTCTTGGATCAGAGTGAG	600
DB	541	CATCTGTAAGAGGACGCTGGTGGCCCGCAGAGATCAACATTTCTTGGATCAGAGTGAG	600
QY	601	GGGTCCCTGCTGTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGAT	660
DB	601	GGGTCCCTGCTGTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGAT	660
QY	661	GACATCGAGACCCCGCATGAGACCGACTCGCTGGAGTTCTTGGGGAATGGCAACGAACTG	720
DB	661	GACATCGAGACCCCGCATGAGACCGACTCGCTGGAGTTCTTGGGGAATGGCAACGAACTG	720
QY	721	GAGTGGGAGACGACACCCCGTGGCCACCGCCAGAGACATGCCCGGGGACAGCGCGAT	780
DB	721	GAGTGGGAGACGACACCCCGTGGCCACCGCCAGAGACATGCCCGGGGACAGCGCGAT	780
QY	781	CTATTGGGGAGCGCACCGAGGAGCGGCGCGCCGCAACGCGCGCTGTGGGCGGACA	840
DB	781	CTATTGGGGAGCGCACCGAGGAGCGGCGCGCCGCAACGCGCGCTGTGGGCGGACA	840
QY	841	GTGATATCGGGGAGCAAGAGCA CGGTATAGACTTGCACATGATCCGGCTTTACATGAAA	900
DB	841	GTGATATCGGGGAGCAAGAGCA CGGTATAGACTTGCACATGATCCGGCTTTACATGAAA	900
QY	901	GTGTGACCCCAAGGAGGTACTACCGGCAAGGCTCAACGCCCATCATCGTCTTCCAGCC	960
DB	901	GTGTGACCCCAAGGAGGTACTACCGGCAAGGCTCAACGCCCATCATCGTCTTCCAGCC	960
QY	961	TGCTTCTCTTCCAGACAGAGCGCTCCCGACTACCACTACATCATCGAGAGAACTCTTCTG	1020
DB	961	TGCTTCTCTTCCAGACAGAGCGCTCCCGACTACCACTACATCATCGAGAGAACTCTTCTG	1020
QY	1021	TACGTATCAGCAGCTTACAGCTCTTGGTGGTGGAGGACTACATGATCGTGTACCTGAAAC	1080
DB	1021	TACGTATCAGCAGCTTACAGCTCTTGGTGGTGGAGGACTACATGATCGTGTACCTGAAAC	1080
QY	1081	GSTGCCACGCGCGGGGAGGATGCTTGAATCGGCTGGCTGAAGAGTGTACCAAGATG	1140
DB	1081	GSTGCCACGCGCGGGGAGGATGCTTGAATCGGCTGGCTGAAGAGTGTACCAAGATG	1140
QY	1141	ATCGACCGGAGGTTCGGGAAAAAAGCTGAAAGTCTTGTGATCATCGTCCACCCCTCGTGTTC	1200
DB	1141	ATCGACCGGAGGTTCGGGAAAAAAGCTGAAAGTCTTGTGATCATCGTCCACCCCTCGTGTTC	1200
QY	1201	ATTGCGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGGCTCAAGTTCATCAACAGATC	1260
DB	1201	ATTGCGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGGCTCAAGTTCATCAACAGATC	1260
QY	1261	CAGTACGTGCACAGCTTGGAGACCTGGAGCAACTCATCCCTATGGAAACCGTCCAGATC	1320
DB	1261	CAGTACGTGCACAGCTTGGAGACCTGGAGCAACTCATCCCTATGGAAACCGTCCAGATC	1320

QY	1321	CCAGACTGGTCTGCTCAATACGAGAGGAAAGCTGAGGCCGAGGAGAGCGGAGG	1380
Db	1321	CCAGACTGGTCTGCTCAATACGAGAGGAAAGCTGAGGCCGAGGAGAGCGGAGG	1380
QY	1381	CCCCAGCCGAGGTTTGTGCTGCCAGCTCTGAGAGAGGCGAGGTTGGACACGATGGAA	1440
Db	1381	CCCCAGCCGAGGTTTGTGCTGCCAGCTCTGAGAGAGGCGAGGTTGGACACGATGGAA	1440
QY	1441	AACAGGTCTGCTCTGGTCTCAGAGATCAGGAAACAGAGCATGCTCTGAGGCGACGTGAGC	1500
Db	1441	AACAGGTCTGCTCTGGTCTCAGAGATCAGGAAACAGAGCATGCTCTGAGGCGACGTGAGC	1500
QY	1501	ATAACAAAGGACATGGAAGAGATTCAGATGCGAGAAACCTCTGTCTAGACGCCCACTG	1560
Db	1501	ATAACAAAGGACATGGAAGAGATTCAGATGCGAGAAACCTCTGTCTAGACGCCCACTG	1560
QY	1561	GCCCCAGATCTCATCTGCTCTCATCTGAGTCCCAATCTTCCAGGGTCCAGGCCCTCC	1620
Db	1561	GCCCCAGATCTCATCTGCTCTCATCTGAGTCCCAATCTTCCAGGGTCCAGGCCCTCC	1620
QY	1621	GTTCATCTCTGAAACCCAGCATCTTTTCAGCTGCTTGAAGAACATTTGATTTTTTTTTT	1680
Db	1621	GTTCATCTCTGAAACCCAGCATCTTTTCAGCTGCTTGAAGAACATTTGATTTTTTTTTT	1680
QY	1681	TAACGATCAGTATTTGTGGTTCAGAAAGGCGCCAGCTCTGAGCCCTCACCTTCC	1740
Db	1681	TAACGATCAGTATTTGTGGTTCAGAAAGGCGCCAGCTCTGAGCCCTCACCTTCC	1740
QY	1741	ACACTCAGCAACTCTCAGCCGAGGAGGCAAGGCGCAGGGGTGGCCGGTGGCGTC	1800
Db	1741	ACACTCAGCAACTCTCAGCCGAGGAGGCAAGGCGCAGGGGTGGCCGGTGGCGTC	1800
QY	1801	GGTGGCTCCGCTCTGCTCGCAGCCCTCTGCTGAGCTGGATACAGATTCAGAGCC	1860
Db	1801	GGTGGCTCCGCTCTGCTCGCAGCCCTCTGCTGAGCTGGATACAGATTCAGAGCC	1860
QY	1861	CTTCTCTGCTGCTCACCCTCAGGTTGGAGGCAAGACCCCGCCACCCCGCT	1920
Db	1861	CTTCTCTGCTGCTCACCCTCAGGTTGGAGGCAAGACCCCGCCACCCCGCT	1920
QY	1921	GGGTCTGCTGCTTCTGCTGCTTCCCTCCAGATGCGGCTCAGACTAGAGCTCA	1980
Db	1921	GGGTCTGCTGCTTCTGCTGCTTCCCTCCAGATGCGGCTCAGACTAGAGCTCA	1980
QY	1981	ACCCCTCTATGAGGCGCAGCTCTGCGGTAGCTCTGACCTCCGACCTTATGTCGAAT	2040
Db	1981	ACCCCTCTATGAGGCGCAGCTCTGCGGTAGCTCTGACCTCCGACCTTATGTCGAAT	2040
QY	2041	TCACACCCATGTTTTCATTTGACCCGCCCTTCTCGCTCATATGACACCCAGCTCC	2100
Db	2041	TCACACCCATGTTTTCATTTGACCCGCCCTTCTCGCTCATATGACACCCAGCTCC	2100
QY	2101	TTTGAGAGGATCAGAGCCCATTCAGAGAGAGCGCTGCCAACCATCTTGTCTCCG	2160
Db	2101	TTTGAGAGGATCAGAGCCCATTCAGAGAGAGCGCTGCCAACCATCTTGTCTCCG	2160
QY	2161	ATTGCAAAATGACACCCAGTAACTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCA	2220
Db	2161	ATTGCAAAATGACACCCAGTAACTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCA	2220
QY	2221	AGCCACCGGCAACCCCGTCAATACCTCCCAAGGAATGAGATATGTGACCTCACT	2280
Db	2221	AGCCACCGGCAACCCCGTCAATACCTCCCAAGGAATGAGATATGTGACCTCACT	2280
QY	2281	GCTCCCCCAACCCAGGCTGAGGTCAGCAGCCCAACGCTGTTCCGGTTGGAACAGCAG	2340
Db	2281	GCTCCCCCAACCCAGGCTGAGGTCAGCAGCCCAACGCTGTTCCGGTTGGAACAGCAG	2340
QY	2341	AGGCTCAGAAACTGGCTCTGAAATAGGCAAGCTTAGCAGAGGAAGATACAGGGTATCGG	2400
Db	2341	AGGCTCAGAAACTGGCTCTGAAATAGGCAAGCTTAGCAGAGGAAGATACAGGGTATCGG	2400
QY	2401	CGGTTTGTAGTGTTCAGAAAGTCAATTCGGGAAGATAAATCCAGTGGCTGGCGGAGCCAC	2460

Db	2401																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Matches 2911; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy	1	GCCGAGCCTCTGACAGCCCTGAGCTGGGAAGACAGCTACCTCGGAGGACAGGGCGCGCA	60
Db	5	GCCGAGCCTCTGACAGCCCTGAGCTGGGAAGACAGCTACCTCGGAGGACAGGGCGCGCA	64
Qy	61	GGCGGGGGGGATGAGAGGGGGGGAGCCGCGAGCCCGCGCTGGGGAGCCCAACCGCTAAC	120
Db	65	GGCGGGGGGGATGAGAGGGGGGGAGCCGCGAGCCCGCGCTGGGGAGCCCAACCGCTAAC	124
Qy	121	CTGCAACCCACCCCTGCAACAAAGAGCTGGGGGGCTGGCGAGCTGGCCCTGGG	180
Db	125	CTGCAACCCACCCCTGCAACAAAGAGCTGGGGGGCTGGCGAGCTGGCCCTGGG	184
Qy	181	TGACCTTCTCGGATCGAGAATCCGCCCTTGGGAGCATCTCTTCTCTAGGCTCTGAA	240
Db	185	TGACCTTCTCGGATCGAGAATCCGCCCTTGGGAGCATCTCTTCTCTAGGCTCTGAA	244
Qy	241	GGCCGGGGAGGTGAGGATGCCCAGCTGCACCCGGGGAGGGCTCGCCTTTGTTGGCA	300
Db	245	GGCCGGGGAGGTGAGGATGCCCAGCTGCACCCGGGGAGGGCTCGCCTTTGTTGGCA	304
Qy	301	GTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGTCTATCCCTGCTTCAAGCCAGTGCCTC	360
Db	305	GTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGTCTATCCCTGCTTCAAGCCAGTGCCTC	364
Qy	361	TTCCAGCTCCCATGTTGGGACCAACCGAAGCCACGCTCCGGATGGAAACCGTGGACGTGAAG	420
Db	365	TTCCAGCTCCCATGTTGGGACCAACCGAAGCCACGCTCCGGATGGAAACCGTGGACGTGAAG	424
Qy	421	GAGGAATGGAGGACGAAGATCTTCCAGGCGCATCCCAAGAGAGAGAGCGGGGTGGAACTG	480
Db	425	GAGGAATGGAGGACGAAGATCTTCCAGGCGCATCCCAAGAGAGAGAGCGGGGTGGAACTG	484
Qy	481	CTTGGCAGCCCGTGGAGACACATCTCTCTCCCAACACGCTTAAATTTCAACGGAGCG	540
Db	485	CTTGGCAGCCCGTGGAGACACATCTCTCTCCCAACACGCTTAAATTTCAACGGAGCG	544
Qy	541	CATCGTAAGAGGAAGACGCTGTGGCCCGCAGAGATCAACATTTCTCGATCAGAGTGAG	600
Db	545	CATCGTAAGAGGAAGACGCTGTGGCCCGCAGAGATCAACATTTCTCGATCAGAGTGAG	604
Qy	601	GGTCCCTGTGTCGATGATCTTTGGATACCCCTGATGACTGGATATTAACGTGGAT	660
Db	605	GGTCCCTGTGTCGATGATCTTTGGATACCCCTGATGACTGGATATTAACGTGGAT	664
Qy	661	GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAATGGCAACGAACCTG	720
Db	665	GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAATGGCAACGAACCTG	724
Qy	721	GAGTGGGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCGGGACAGCGCGAT	780
Db	725	GAGTGGGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCGGGACAGCGCGAT	784
Qy	781	CTATTTGGGAGCGGACGAGAGGACGGAGCGCGCCCAACGGGGCGCTGTGGGGAGCA	840
Db	785	CTATTTGGGAGCGGACGAGAGGACGGAGCGCGCCCAACGGGGCGCTGTGGGGAGCA	844
Qy	841	GTGATCATCGGGAGCAAGAGCACCGTATAGACCTGCACATGATCCGGCTTACATGAAA	900
Db	845	GTGATCATCGGGAGCAAGAGCACCGTATAGACCTGCACATGATCCGGCTTACATGAAA	904
Qy	901	GTGGTCAACCCACGAGGGTACTACGGCGAAGGCTTCAACCGCATCATCTCTTCGAGCC	960
Db	905	GTGGTCAACCCACGAGGGTACTACGGCGAAGGCTTCAACCGCATCATCTCTTCGAGCC	964
Qy	961	TGCTTCTCTCCAGACGAGCCTCCCGACTACCACTACATCATGAGAACTCTTTCCTG	1020
Db	965	TGCTTCTCTCCAGACGAGCCTCCCGACTACCACTACATCATGAGAACTCTTTCCTG	1024
Qy	1021	TAGCTCATGAGGCTTACGCTCTGCTGGCTGAGGACTACATGATCGTGTACCTGAC	1080
Db	1025	TAGCTCATGAGGCTTACGCTCTGCTGGCTGAGGACTACATGATCGTGTACCTGAC	1084

Qy	1081	GGTCCACGCCCCCGGAGGATGCTGGAATCGGCTGGCTGAAGAGTGCTTACCAGATG	1140
Db	1085	GGTCCACGCCCCCGGAGGATGCTGGAATCGGCTGGCTGAAGAGTGCTTACCAGATG	1144
Qy	1141	ATCGACCGGAGGTTGCGGAAAAAAGCTGAAAGTCTTTGATCATCGTCCACCCCTCGTGGTTC	1200
Db	1145	ATCGACCGGAGGTTGCGGAAAAAAGCTGAAAGTCTTTGATCATCGTCCACCCCTCGTGGTTC	1204
Qy	1201	ATTGGGACTGTGCTGGCCCATCTCTGCCCTTTTCATCAGCGCTCAAGTTTCATCAACAGATC	1260
Db	1205	ATTGGGACTGTGCTGGCCCATCTCTGCCCTTTTCATCAGCGCTCAAGTTTCATCAACAGATC	1264
Qy	1261	CAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCTTATGGAACAGTCCAGATC	1320
Db	1265	CAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCTTATGGAACAGTCCAGATC	1324
Qy	1321	CCAGACTCGCTCTGCAATACGAAGAGAAAGACTGAAGGCCAGGAGGAGGAGCGGAGG	1380
Db	1325	CCAGACTCGCTCTGCAATACGAAGAGAAAGACTGAAGGCCAGGAGGAGGAGCGGAGG	1384
Qy	1381	CCCCAGCCGAGTTTGTCTGCCAGGTCTGAAGAGAGCCAGAGGTGGCACCACTGGAA	1440
Db	1385	CCCCAGCCGAGTTTGTCTGCCAGGTCTGAAGAGAGCCAGAGGTGGCACCACTGGAA	1444
Qy	1441	AACAGGTCTGCTCTGAGTCTCAGAAAGATCAGGAAACCAAGCATGTCTGAGGGGACCTGAGC	1500
Db	1445	AACAGGTCTGCTCTGAGTCTCAGAAAGATCAGGAAACCAAGCATGTCTGAGGGGACCTGAGC	1504
Qy	1501	ATAACAAAGGACATGGAAGAAAGATTCAGATGCGCAAGAAAACCTCTGTGACAGCCCACTG	1560
Db	1505	ATAACAAAGGACATGGAAGAAAGATTCAGATGCGCAAGAAAACCTCTGTGACAGCCCACTG	1564
Qy	1561	GCCCGAGATCATCTCTGCTCATCTGAGTCCCAATCTTCCAGGGTGGCCAGCCCTCC	1620
Db	1565	GCCCGAGATCATCTCTGCTCATCTGAGTCCCAATCTTCCAGGGTGGCCAGCCCTCC	1624
Qy	1621	GTTCATCTCTGAAACCCAGCATCTCTTTTCAGCTGTCTGAAACATTTGATTTTTTTTTT	1680
Db	1625	GTTCATCTCTGAAACCCAGCATCTCTTTTCAGCTGTCTGAAACATTTGATTTT	1682
Qy	1681	TAAAGATGAGATATTTGTGCGTTTCAGAAAAGGGCCAGCTCTGAGCCCTTCAACCTTCC	1740
Db	1683	TAAAGATGAGATATTTGTGCGTTTCAGAAAAGGGCCAGCTCTGAGCCCTTCAACCTTCC	1742
Qy	1741	ACACTCAGAACTCTCAGCCGAGGAGGACAGAGCGAGGGGTGGCCCGGTGGCGTC	1800
Db	1743	ACACTCAGAACTCTCAGCCGAGGAGGACAGAGCGAGGGGTGGCCCGGTGGCGTC	1802
Qy	1801	GGTGGCTTCCGCTCTCTGCTCGAGCCCTGTGCTGAGAGTGGATACAGATTTCAAGACC	1860
Db	1803	GGTGGCTTCCGCTCTCTGCTCGAGCCCTGTGCTGAGAGTGGATACAGATTTCAAGACC	1862
Qy	1861	CTTCTCTTGTGTCAACCCGCTCAGGTTGAGGACACAGACCCACCGCCACCCCGGT	1920
Db	1863	CTTCTCTTGTGTCAACCCGCTCAGGTTGAGGACACAGACCCACCGCCACCCCGGT	1922
Qy	1921	GGGTCTGCTCTCTTCTGCTGCTTCCCTCCAGAAAGGCGCTCAGACCTAGAGACTCA	1980
Db	1923	GGGTCTGCTCTCTTCTGCTGCTTCCCTCCAGAAAGGCGCTCAGACCTAGAGACTCA	1982
Qy	1981	ACCCCTCTATGAGGGCCACGCTCTGGGGTAGTCTCTGACCTCCGACCTTATGTCCAAAT	2040
Db	1983	ACCCCTCTATGAGGGCCACGCTCTGGGGTAGTCTCTGACCTCCGACCTTATGTCCAAAT	2042
Qy	2041	TCACACCATGTTTTCATTTGACCCCGCTTCTCGCTCATTAATGACACCCAGCTCC	2100
Db	2043	TCACACCATGTTTTCATTTGACCCCGCTTCTCGCTCATTAATGACACCCAGCTCC	2102
Qy	2101	TTTGAGGAGTACAGACCCATTTGCAAGAGAGCGCTGCGCAACCTTGTCTCTCG	2160
Db	2103	TTTGAGGAGTACAGACCCATTTGCAAGAGAGCGCTGCGCAACCTTGTCTCTCG	2162

```
QY 2161 ATTGCAAAATGACACCCAGTATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGCA 2220
Db 2163 ATTGCAAAATGACACCCAGTATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGCA 2222
QY 2221 AGCCACCGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGGACCTCACT 2280
Db 2223 AGCCACCGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGGACCTCACT 2282
QY 2281 GTTCCCTCCCAACCCAGCGTCAAGCTGGGACACGCGCAACGCTGTTCCGGTGGGAACAGCAG 2340
Db 2283 GTTCCCTCCCAACCCAGCGTCAAGCTGGGACACGCGCAACGCTGTTCCGGTGGGAACAGCAG 2342
QY 2341 AGCTCAGAACTGCTCTGAATAGCAGACCTAGCAGAGGAAGATACAGGATATCGG 2400
Db 2343 AGCTCAGAACTGCTCTGAATAGCAGACCTAGCAGAGGAAGATACAGGATATCGG 2402
QY 2401 GCGTTTGAGTGTTCAGAACTCATTCCGGAAGATAAATCAAGTGCCTGGCCGAGCCAC 2460
Db 2403 GCGTTTGAGTGTTCAGAACTCATTCCGGAAGATAAATCAAGTGCCTGGCCGAGCCAC 2462
QY 2461 CTGCAATTCAGAGCTTGGACAGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAAA 2520
Db 2463 CTGCAATTCAGAGCTTGGACAGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAAA 2522
QY 2521 GAAGACAGACTTTTCTAATGTGTCCAAATGCGATCACTGGTCAAGTGAATCTAGAG 2580
Db 2523 GAAGACAGACTTTTCTAATGTGTCCAAATGCGATCACTGGTCAAGTGAATCTAGAG 2582
QY 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGCCAGGCAAGTGGCTCA 2640
Db 2583 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGCCAGGCAAGTGGCTCA 2642
QY 2641 CGCTCTGTAATCCAGACTTTGGGAGGCGGAGGCGAGCGGATCACTGAGTCAAGAGTTT 2700
Db 2643 CGCTCTGTAATCCAGACTTTGGGAGGCGGAGGCGGATCACTGAGTCAAGAGTTT 2702
QY 2701 GAGAACAGCTGGCCAACTGAGTGAACCTCTATCTCTACTAAATAATACAAAAATAGCCA 2760
Db 2703 GAGAACAGCTGGCCAACTGAGTGAACCTCTATCTCTACTAAATAATACAAAAATAGCCA 2762
QY 2761 GCGTGTGGCAGGTCCTGTAATCCAGTACTTGGGAGGCTGGGAGGCTGAGGATGAGATCACT 2820
Db 2763 GCGTGTGGCAGGTCCTGTAATCCAGTACTTGGGAGGCTGGGAGGCTGAGGATGAGATCACT 2822
QY 2821 TAAACCTGAGAGGAGGCTTACAGTGAAGCAAGATCGTGCCACTGCACTTCAGAGCTGGG 2880
Db 2823 TAAACCTGAGAGGAGGCTTACAGTGAAGCAAGATCGTGCCACTGCACTTCAGAGCTGGG 2882
QY 2881 CGACAGAGCAAGACTCTGTCTCAAAAAAATAAAAAA 2917
Db 2883 CGACAGAGCAAGACTCTGTCTCAAAAAAATAAAAAA 2919

RESULT 3
LOCUS BD183464
DEFINITION Novel genes and proteins encoded by the genes.
ACCESSION BD183464
VERSION BD183464.1 GI:31875664
KEYWORDS JP 2002345492-A/177.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 4976)
REFERENCE
AUTHORS Ohara, O., Nagase, T. and Nakajima, D.
TITLE Novel genes and proteins encoded by the genes.
JOURNAL Patent: JP 2002345492-A 177 03-DEC-2002;
KAZUSA DNA RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002345492-A/177
PD 03-DEC-2002
```

```
PF 26-FEB-2002 JP 2002049009
PI OSAMU OHARA, TAKAHIRO NAGASE, DAISUKE NAKAJIMA
PC C12N15/09, C07K14/47//A61K31/711, A61K38/00, A61K48/00, A61P25/00,
PC A61P25/14,
PC A61P25/18, A61P35/00, C12N15/00, A61K37/02
CC Novel genes and proteins encoded by the genes FH Key
FT CDS Location/Qualifiers
(404)..(1489).
FEATURES
Location/Qualifiers
source 1..4976
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.3%; Score 2896.6; DB 6; Length 4976;
Best Local Similarity 99.8%; Fred. No. 0;
Matches 2911; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
QY 1 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGAGCAGCTACCTCGGAGGACGGCGCGCA 60
Db 5 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGAGCAGCTACCTCGGAGGACGGCGCGCA 64
QY 61 GCGCGGCGGCGATGAGAGGGGGCGCAGCGCAGCCCGCGCTGGGGAGGCCACCGCTAAC 120
Db 65 GCGCGGCGGCGATGAGAGGGGGCGCAGCGCAGCCCGCGCTGGGGAGGCCACCGCTAAC 124
QY 121 CTGCAACCCACCCACCCCTGCAACAAAGAGCTGGGGGGCGTGGCCAGCTGGCCCTGGG 180
Db 125 CTGCAACCCACCCACCCCTGCAACAAAGAGCTGGGGGGCGTGGCCAGCTGGCCCTGGG 184
QY 181 TGACCTTCTCGATGTCAGAAATCCGCCCTCGGAGCATCTCTTCTCTAGGCTCTGAA 240
Db 185 TGACCTTCTCGATGTCAGAAATCCGCCCTCGGAGCATCTCTTCTCTAGGCTCTGAA 244
QY 241 GCGCGGAGGCGTGGAGCGATGCCAGCTGCAACCGGCGAGGGCTCGCTTGTGTTGCCA 300
Db 245 GCGCGGAGGCGTGGAGCGATGCCAGCTGCAACCGGCGAGGGCTCGCTTGTGTTGCCA 304
QY 301 GTAGGAGAGAGGCTGTCTGAGCTGAGAGGGGTTCCTCTTCAAGCAGATGCTCTC 360
Db 305 GTAGGAGAGAGGCTGTCTGAGCTGAGAGGGGTTCATCTCTTCAAGCAGATGCTCTC 364
QY 361 TTCCAGCTCCCATGGGACCAACGAGCCAGCTCCCGATGGAAGAGAGCGGGGTGGAAG 420
Db 365 TTCCAGCTCCCATGGGACCAACGAGCCAGCTCCCGATGGAAGAGAGCGGGGTGGAAG 424
QY 421 GAGGAATGGCAGGACGAAAGATCTTCCAGGGCCACTCCCGAGAGAGAGCGGGGTGGAAG 480
Db 425 GAGGAATGGCAGGACGAAAGATCTTCCAGGGCCACTCCCGAGAGAGAGCGGGGTGGAAG 484
QY 481 CTGGCAGCCCGGTGGAAAGACATCTCTCTCTCCCAACAGCTAAATTTCAACCGAGCG 540
Db 485 CTGGCAGCCCGGTGGAAAGACATCTCTCTCTCCCAACAGCTAAATTTCAACCGAGCG 544
QY 541 CATCTAGAGAGAGAGCGCTGGTGGCCCCCAGAGATCAACATTTCTCTGATCAGAGTGA 600
Db 545 CATCTAGAGAGAGAGCGCTGGTGGCCCCCAGAGATCAACATTTCTCTGATCAGAGTGA 604
QY 601 GGGTCCCTGCTGTCCGATGACTTTTGGATATCCCTGATGACCTGGATATTAACCTGGAT 660
Db 605 GGGTCCCTGCTGTCCGATGACTTTTGGATATCCCTGATGACCTGGATATTAACCTGGAT 664
QY 661 GACATCGAGAGACCCCGATGAGACCGGACTCGCTGGAGTTCTCTGGGAAATGGCAACGAT 720
Db 665 GACATCGAGAGACCCCGATGAGACCGGACTCGCTGGAGTTCTCTGGGAAATGGCAACGAT 724
QY 721 GAGTGGGAGAGAGACACCCCGTGGCCAGCCCAAGAAATGCTCCCGGGGAGACGCGCGAT 780
Db 725 GAGTGGGAGAGAGACACCCCGTGGCCAGCCCAAGAAATGCTCCCGGGGAGACGCGCGAT 784
QY 781 CTATTGTTGGGAGCGGACGAGCGAGCGCGCCCAACGGCGCTCTGTGGCGGACA 840
```

Db 785 CTATTTGGGACGGCAGCAGGAGGAGCGGCGCGCCGCAACGGGGCCCTGTGGCGGACA 844
Qy 841 GTGATCATCGGGAGCAAGAGCACCCTATAGACCTGCAACATGATCGGGCTTACATGAAA 900
Db 845 GTGATCATCGGGAGCAAGAGCACCCTATAGACCTGCAACATGATCGGGCTTACATGAAA 904
Qy 901 GTGGTCACCCACGGAGGCTACTAGGGGGAAGGCTCAACGGCATCATGCTCTTCGAGCC 960
Db 905 GTGGTCACCCACGGAGGCTACTAGGGGGAAGGCTCAACGGCATCATGCTCTTCGAGCC 964
Qy 961 TGTCTCTCCACAGACAGCCTCCCGACTACCATCATCATGAGAACCTCTTCCTG 1020
Db 965 TGTCTCTCCACAGACAGCCTCCCGACTACCATCATCATGAGAACCTCTTCCTG 1024
Qy 1021 TACGTATCATCAGCAGCTTAGAGCTCTCGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1080
Db 1025 TACGTATCATCAGCAGCTTAGAGCTCTCGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1084
Qy 1081 GGTGCCACGCCCGGGAGGATGCCTGGAACTGGCTGGCTGAAGAGTGCTACAGATG 1140
Db 1085 GGTGCCACGCCCGGGAGGATGCCTGGAACTGGCTGGCTGAAGAGTGCTACAGATG 1144
Qy 1141 ATCGACCGGAGTTGCGGAAAACCTGAAGTCTCTCATCATGTCACCCCTCGTGTTC 1200
Db 1145 ATCGACCGGAGTTGCGGAAAACCTGAAGTCTCTCATCATGTCACCCCTCGTGTTC 1204
Qy 1201 ATTGGACTGTGTGGCCATCTCTCGCCCTTTTCATCAGGGTCAAGTTTCATCAACAGATC 1260
Db 1205 ATTGGACTGTGTGGCCATCTCTCGCCCTTTTCATCAGGGTCAAGTTTCATCAACAGATC 1264
Qy 1261 CAGTACGTGCAAGCTTTGGAAGACCTGGAGCAACTCATCTCTATGAAACAGTCCAGATC 1320
Db 1265 CAGTACGTGCAAGCTTTGGAAGACCTGGAGCAACTCATCTCTATGAAACAGTCCAGATC 1324
Qy 1321 CCAGATCGGTCTGCAATACGAAGAGGAAAGTCAAGGCTGAAAGGAGGAGGAGCGGAG 1380
Db 1325 CCAGATCGGTCTGCAATACGAAGAGGAAAGTCAAGGCTGAAAGGAGGAGGAGCGGAG 1384
Qy 1381 CCCAGCCGGAGTTGTGTGCTGCCAGGTCTGAAGAGAAGCCAGAGGTGGCACAGTGGAA 1440
Db 1385 CCCAGCCGGAGTTGTGTGCTGCCAGGTCTGAAGAGAAGCCAGAGGTGGCACAGTGGAA 1444
Qy 1441 AACAGTCTGTCTGTCTCAGAAATCAGGAAACAGCATGTCTGAGGGGAGCGTGAAGC 1500
Db 1445 AACAGTCTGTCTGTCTCAGAAATCAGGAAACAGCATGTCTGAGGGGAGCGTGAAGC 1504
Qy 1501 ATAAACAGGACATGAAAGATCCAGATGCCAGAAACCTCTGTCAGAGCCCACTG 1560
Db 1505 ATAAACAGGACATGAAAGATCCAGATGCCAGAAACCTCTGTCAGAGCCCACTG 1564
Qy 1561 GCCCAGATCTCATCTCTGCTCATCTCTGATGCCAATCTTCAAGGGTGGCAGCCCTCC 1620
Db 1565 GCCCAGATCTCATCTCTGCTCATCTCTGATGCCAATCTTCAAGGGTGGCAGCCCTCC 1624
Qy 1621 GTTCATCTCGAAACCCAGATCCTTTTCAGTGTCTGAAACATTTGTTTTTTTT 1680
Db 1625 GTTCATCTCGAAACCCAGATCCTTTTCAGTGTCTGAAACATTTGTTTTTTTT 1682
Qy 1681 TAACGATGAGATTTGTGGTTCCAGAAAGGGCCAGCTCTGAGCCCTCACCCCTCC 1740
Db 1683 TAACGATGAGATTTGTGGTTCCAGAAAGGGCCAGCTCTGAGCCCTCACCCCTCC 1742
Qy 1741 ACACTACGAACTCTCAGCGGAGGAGGCAAGAGCGCAGGGGTGGCCCGGTGGCGTC 1800
Db 1743 ACACTACGAACTCTCAGCGGAGGAGGCAAGAGCGCAGGGGTGGCCCGGTGGCGTC 1802
Qy 1801 GGTGGCTCCGCTCTGTCTGAGCCCTCTGTGTGATGATCAAGATTCAGAGCC 1860
Db 1803 GGTGGCTCCGCTCTGTCTGAGCCCTCTGTGTGATGATCAAGATTCAGAGCC 1862
Qy 1861 CTTCCTCTGTCTGATCCCGCTCAGGTTGGAGCCACAGACCCACCGCCCGGCT 1920
Db 1863 CTTCCTCTGTCTGATCCCGCTCAGGTTGGAGCCACAGACCCACCGCCCGGCT 1922

RESULT 4
AB058775
LOCUS

AB058775

4976 bp

mRNA

linear

PRI 05-JUN-2001

Qy 1921 GGGCTGTGGTCTCTTCTGTGCTTTCCCTCCAGAAATCGGCTCAGACCTAGAAAGCTCA 1980
Db 1923 GGGCTGTGCTCTCTTCTGTGCTTTCCCTCCAGAAATCGGCTCAGACCTAGAAAGCTCA 1982
Qy 1981 ACCCCCTATGAGGGCCACGCTCTGGGTAGTCTCTGACCTCCGACCTTATGTCCAAAT 2040
Db 1983 ACCCCCTATGAGGGCCACGCTCTGGGTAGTCTCTGACCTCCGACCTTATGTCCAAAT 2042
Qy 2041 TCACACCCCATGTTTTTTCATTTGACCCGCCCTTCTCGCTCATATATGACCCAGCTCC 2100
Db 2043 TCACACCCCATGTTTTTTCATTTGACCCGCCCTTCTCGCTCATATATGACCCAGCTCC 2102
Qy 2101 TTTGAGAGGATCAGAGCCCATTTGCAACAAGAGAGCGCTGCCAACCATCTTGTCTCG 2160
Db 2103 TTTGAGAGGATCAGAGCCCATTTGCAACAAGAGAGCGCTGCCAACCATCTTGTCTCG 2162
Qy 2161 ATTGCAAAATGACACCCCGAGTAATCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGCA 2220
Db 2163 ATTGCAAAATGACACCCCGAGTAATCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGCA 2222
Qy 2221 AGCCACCGGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGACCTCACT 2280
Db 2223 AGCCACCGGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGACCTCACT 2282
Qy 2281 GCTCCCCCAACCCAGCGTCAGGCTGGGACACGCGCAACGCTGTCGGGTGGAAACAGCAG 2340
Db 2283 GCTCCCCCAACCCAGCGTCAGGCTGGGACATGCGCAACGCTGTCGGGTGGAAACAGCAG 2342
Qy 2341 AGGCTCAGAAACTGGCTCTGAAATAGGACAGCTTAGCAAGAGGAAGATACAGGGTATCGG 2400
Db 2343 AGGCTCAGAAACTGGCTCGGAAATAGGACAGCTTAGCAAGAGGAAGATACAGGGTATCGG 2402
Qy 2401 GGGTTGAGTCTTTTCAGAAAGTCAATTCGGGAAGATAAATCCAGTGGCTGGCGGAGCCAC 2460
Db 2403 GGGTTGAGTCTTTTCAGAAAGTCAATTCGGGAAGATAAATCCAGTGGCTGGCGGAGCCAC 2462
Qy 2461 CTGCATTCAGAACTGGACCGGGTTCTTTGTCGGGAGGCAAAATTTCCCTAGGAAAA 2520
Db 2463 CTGCATTCAGAACTGGACCGGGTTCTTTGTCGGGAGGCAAAATTTCCCTAGGAAAA 2522
Qy 2521 GAAGACAGATTTTCTAATGTGGTCCAAATGCGGATCACTGGTCAAGATGGACTCTAGAAG 2580
Db 2523 GAAGACAGATTTTCTAATGTGGTCCAAATGCGGATCACTGGTCAAGATGGACTCTAGAAG 2582
Qy 2581 CACTGAGCTCCCTGTCTCTGGAATTTTAAAGAAAGGCTGGCCAGGACGATGGCTCA 2640
Db 2583 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAAGAAAGGCTGGCCAGGACGATGGCTCA 2642
Qy 2641 CGCCTGTAAATCCAGACTTTTGGGAGCGGAGCGAGCGGATCACCTGAGGTGAGGAGTTT 2700
Db 2643 CGCCTGTAAATCCAGACTTTTGGGAGCGGAGCGGAGCGGATCACCTGAGGTGAGGAGTTT 2702
Qy 2701 GAGAACAGCTGGCCAAACATGTTGAAACCTCATCTCTACTAATAAATACAAAATTTAGCCA 2760
Db 2703 GAGAACAGCTGGCCAAACATGTTGAAACCTCATCTCTACTAATAAATACAAAATTTAGCCA 2762
Qy 2761 GGGCTGTGGGAGGCTGCTTAATCCAGCTACTTGGGAGGCTGAGGCAATGAGATCACT 2820
Db 2763 GGGCTGTGGGAGGCTGCTTAATCCAGCTACTTGGGAGGCTGAGGCAATGAGATCACT 2822
Qy 2821 TAAACCTCAGAGGCGAGGTTACAGTGAGCCAAAGATCGTGCCACTGCAATCCAGCTGGG 2880
Db 2823 TAAACCTCAGAGGCGAGGTTACAGTGAGCCAAAGATCGTGCCACTGCAATCCAGCTGGG 2882
Qy 2881 CGACAGACGAGACTCTGTCTCAAAAAAATAAAAAA 2917
Db 2883 CGACAGACGAGACTCTGTCTCAAAAAAATAAAAAA 2919

```

DEFINITION Homo sapiens mRNA for KIAA1872 protein, partial cds.
ACCESSION AB058775
VERSION AB058775.1 GI:14017960
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
MAMMALIA; EUTHERIA; EUARCHONTOGRIRES; PRIMATES; CATARTHINI;
HOMINIDAE; HOMO.
REFERENCE Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R. and Ohara, O.
AUTHORS Prediction of the coding sequences of unidentified human genes. XX.
TITLE The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro
JOURNAL DNA Res. 8 (2), 85-95 (2001)
PUBMED 11347906
REFERENCE 2 (bases 1 to 4976)
AUTHORS Ohara, O., Nagase, T. and Kikuno, R.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,
Fax: 81-438-52-3914)
FEATURES             Location/Qualifiers
     source            1..4976
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="bj05518"
                     /tissue_type="brain"
                     /note="vector: pBluescriptII SK plus"
     gene              1..4976
                     /gene="KIAA1872"
     CDS               <188..1492
                     /gene="KIAA1872"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /product="KIAA1872 protein"
                     /protein_id="BAB47501.1"
                     /db_xref="GI:14017961"
                     /translation="PSSDARSAPASILFLGSEGPSVSDAQLHPRARLCLPVRRRG
                     CLSCGVIPASQCLFPAPMGTEATLRNENVDVKEWQEDLPRPLPEETGVELLGS
                     PVEDSSPTNLFNGAHRKRKTLVAPEINISLDQSEGLSDDFLDTPDDLDINVDI
                     IETPDTSLEFLNGNELEWEDDTPVATKNMGDSADLFGDFTTDSGAANGRLWR
                     TVIIGQEHRIIDLHMRPYMKVTHGGYVGEGLNAILVFAACFLPDSILPDVHYIMEN
                     LFLVYISLELLAVEDYIMIVLNGATPRRMPGICWKLKCYQIMIDRLRKNLXSLIV
                     HPSMRTVLAISRPPISVKFINKIQYVHSLEDLSQLIPMERHVQIPDCVLTVEERLK
                     ARRESARPOPEFLPRSEKPEVAPNRSALVSDQETSMS"
ORIGIN
Query Match          99.3%; Score 2896.6; DB 8; Length 4976;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2911; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 1  GCCGAGCCTCTGCCAGCCTGAGCTGGGAAGAGCAGCTACCTCGGAGGACGGGCGCGCA 60
DB 5  GCCGAGCCTCTGCCAGCCTGAGCTGGGAAGAGCAGCTACCTCGGAGGACGGGCGCGCA 64
QY 61  GCGGGGCGGCGATGAGAGGGGCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 65  GCGGGGCGGCGATGAGAGGGGCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 124
QY 121  CTTGCAACCCACCCACCCCTGCACAAAGAGCTGGGGGCGGCTGGCGGCGGCGGCGGCGG 180
DB 125  CTTGCAACCCACCCACCCCTGCACAAAGAGCTGGGGGCGGCTGGCGGCGGCGGCGGCGG 184
QY 181  TGAACCTTCTCGGATGCAAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 185  TGAACCTTCTCGGATGCAAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 244
QY 241  GCGCGGGGAGCGTGAGCGATGCCAGCTGCACCGGCGGCGGCGGCGGCGGCGGCGGCGG 300

```


QY	1381	CCCACGCGGAGTTTGTGCTGCGCCAGTCTGAAGAGAGCCAGAGTGGCACAGTGGAA	1440
Db	1385	CCCACGCGGAGTTTGTGCTGCGCCAGTCTGAAGAGAGCCAGAGTGGCACAGTGGAA	1444
QY	1441	AACAGTCTGCTCTGCTCTCAGAGATCAGGAACAAGCATGTCTGAGGCGACGTGAGC	1500
Db	1445	AACAGTCTGCTCTGCTCTCAGAGATCAGGAACAAGCATGTCTGAGGCGACGTGAGC	1504
QY	1501	ATAACAAAGACATGGAAGAAGATTCAGATGCCAGAAACCTCTGTCAAGAGCCCACTG	1560
Db	1505	ATAACAAAGACATGGAAGAAGATTCAGATGCCAGAAACCTCTGTCAAGAGCCCACTG	1564
QY	1561	GCCCCAGATCTCATCTGCTGCTCATCTGTAGTCCCAATCTTCCAAAGGTGCCAGCCCTCC	1620
Db	1565	GCCCCAGATCTCATCTGCTGCTCATCTGTAGTCCCAATCTTCCAAAGGTGCCAGCCCTCC	1624
QY	1621	GTTTCATCTCTGAACCCAGCATCTTTTCAGCTGCTTGAAACAATTGTATTTTTTTTTT	1680
Db	1625	GTTTCATCTCTGAACCCAGCATCTTTTCAGCTGCTTGAAACAATTGTATTTTTTTTTT	1682
QY	1681	TAAAGATGAGATTTTGTGCTTCCAGAAAGGCGCCAGCTCTGAGCCCTCACCCCTCC	1740
Db	1683	TAAAGATGAGATTTTGTGCTTCCAGAAAGGCGCCAGCTCTGAGCCCTCACCCCTCC	1742
QY	1741	ACACTCAGAACTCTCAGCCGAGGAAGCAAGAGCGCAGGGGTGGCCGGTGGCGTC	1800
Db	1743	ACACTCAGAACTCTCAGCCGAGGAAGCAAGAGCGCAGGGGTGGCCGGTGGCGTC	1802
QY	1801	GGTGGCTCGCTCTGCTGCGAGCCCTGTGTGTGAGCTGATCAAGATTCAAGACC	1860
Db	1803	GGTGGCTCGCTCTGCTGCGAGCCCTGTGTGTGAGCTGATCAAGATTCAAGACC	1862
QY	1861	CTTCTCTGCTGTCAACCGCTCAGGTTGGAGCCACAGACCCACCGCCCGCGCT	1920
Db	1863	CTTCTCTGCTGTCAACCGCTCAGGTTGGAGCCACAGACCCACCGCCCGCGCT	1922
QY	1921	GGGTCTGCTCTTCTGCTCTTCCCTCCAGAAATGGGCTCAGACCTAGAACTCA	1980
Db	1923	GGGTCTGCTCTTCTGCTCTTCCCTCCAGAAATGGGCTCAGACCTAGAACTCA	1982
QY	1981	ACCCCTTAGAGGCGCATCTCTGGGTAGTCTTGACCTCGACCTTATGTCCAAAT	2040
Db	1983	ACCCCTTAGAGGCGCATCTCTGGGTAGTCTTGACCTCGACCTTATGTCCAAAT	2042
QY	2041	TCACACCATGGTTTTCATTTGACCGCCCTCTCTGCTCATATGACACCCAGCTCC	2100
Db	2043	TCACACCATGGTTTTCATTTGACCGCCCTCTCTGCTCATATGACACCCAGCTCC	2102
QY	2101	TTTGAGAGGATCAGAGCCCATTTGCAAGAAGAGCGCTGCCAACCATCTTGTCTCCG	2160
Db	2103	TTTGAGAGGATCAGAGCCCATTTGCAAGAAGAGCGCTGCCAACCATCTTGTCTCCG	2162
QY	2161	ATTGCAAAATGACACCCAGTATCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGCA	2220
Db	2163	ATTGCAAAATGACACCCAGTATCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGCA	2222
QY	2221	AGCCACCGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGACCTCACT	2280
Db	2223	AGCCACCGGCAAAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGACCTCACT	2282
QY	2281	GCTCCCCCAACCCAGCGTCAGGTGGGACACGCAACGCTGTTCCGGGTTGGAACAGCAG	2340
Db	2283	GCTCCCCCAACCCAGCGTCAGGTGGGACATGCCAACGCTGTTCCGGGTTGGAACAGCAG	2342
QY	2341	AGGCTCAGAACTGCGTCTGAATAGGCGACCTAGCAAGAGAGATACAGGGTATCGG	2400
Db	2343	AGGCTCAGAACTGCGTCTGGAATAGGCGACCTAGCAAGAGAGATACAGGGTATCGG	2402
QY	2401	GCGTTTGTGTTTCAAGAGTCAATTCGGGAAGATTAATCCAGTGGCTGCGCGCAGCCAC	2460
Db	2403	GCGTTTGTGTTTCAAGAGTCAATTCGGGAAGATTAATCCAGTGGCTGCGCGCAGCCAC	2462
QY	2461	CTGCATTCAAAGCTTTGGACCGGCTTTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA	2520
Db	2463	CTGCATTCAAAGCTTTGGACCGGCTTTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA	2522
QY	2521	GAAGACAGACTTTTCTTAATGTGGTCCAAATGCGGATCACTGGTCAAGTGGACTCTAGAG	2580
Db	2523	GAAGACAGACTTTTCTTAATGTGGTCCAAATGCGGATCACTGGTCAAGTGGACTCTAGAG	2582
QY	2581	CACTCAGCTCCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGCCAGGACGATGGCTCA	2640
Db	2583	CACTCAGCTCCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGCCAGGACGATGGCTCA	2642
QY	2641	CGCTGTATCCAGACTTTTGGGAGCGGAGCGGCGGATCACCTGAGGTGAGGATTTT	2700
Db	2643	CGCTGTATCCAGACTTTTGGGAGCGGAGCGGCGGATCACCTGAGGTGAGGATTTT	2702
QY	2701	GAGAACAGCTGGGCCAACATGGTGAACCTCATCTCTAATAAATACAAAAATTAGCCA	2760
Db	2703	GAGAACAGCTGGGCCAACATGGTGAACCTCATCTCTAATAAATACAAAAATTAGCCA	2762
QY	2761	GCGGTGGTGGCAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGATCACT	2820
Db	2763	GCGGTGGTGGCAGGTGCTGTAAATCCAGCTACTTGGGAGGCTGAGGATCACT	2822
QY	2821	TAAACCTGAGAGGCGAGGTTTACAGTGAGCCAAAGATCGTGCCACTTCCAGCTGGG	2880
Db	2823	TAAACCTGAGAGGCGAGGTTTACAGTGAGCCAAAGATCGTGCCACTTCCAGCTGGG	2882
QY	2881	CGACAGAGCAAGACTCTCTCTCAAAAAAATAAAAAA 2917	
Db	2883	CGACAGAGCAAGACTCTCTCTCAAAAAAATAAAAAA 2919	
RESULT 5			
AR560886			
LOCUS			
DEFINITION			
Sequence 1 from patent US 6756212.			
ACCESSION			
AR560886			
VERSION			
AR560886.1 GI:53973364			
KEYWORDS			
Unknown.			
SOURCE			
ORGANISM			
Unknown.			
REFERENCE			
1 (bases 1 to 3076)			
Curtis, R. A. J. and Glucksmann, M. A.			
Isolated proteins and nucleic acid molecules having homology to the			
Nif2 protein and uses thereof			
Patent: US 6756212-A 1 29-JUN-2004;			
Millennium Pharmaceuticals, Inc.; Cambridge, MA;			
EPX;			
FEATURES			
source			
Location/Qualifiers			
1..3076			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Query Match			
Best Local Similarity 99.0%; Score 2887.6; DB 6; Length 3076;			
Matches 2890; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	24	CTGGGAAGAAGCAGCTACCTCGAGGCGAGGCGGCGGCGGCGGCGATGAGAGGGGCG	83
Db	21	CGGGGAAGAAGCAGCTACCTCGAGGCGAGGCGGCGGCGGCGGCGGCGATGAGAGGGGCG	80
QY	84	GCAGCGCAGCGCCCGCGCTGGGGAGCCCAACCGCTAACCTTGACCCCAACCCCTCGA	143
Db	81	GCAGCGCAGCGCCCGCGCTGGGGAGCCCAACCGCTAACCTTGACCCCAACCCCTCGA	140
QY	144	CAAAAGAGCTGGCGGGCGCTGCCACGTCGCCCTGGGGTACCTTCTCGGATGCAAGATC	203
Db	141	CAAAAGAGCTGGCGGGCGCTGCCACGTCGCCCTGGGGTACCTTCTCGGATGCAAGATC	200
QY	204	CGCCCTCGGAGCATCTCTTCTCTTAGGCTCTGAAGGCCCGGGGAGCGTGGAGCGATGC	263

Db 201 |||||CGCCCCGCGAGATCTCTCTCTCTAGGCTCTGAAGGCCCGGGGAGCGGTGAGCGATGC 260
QY 264 CCAAGTGCACCCCGGCGAGGCTCGCCCTTTGTTTGCAGTAAGAGGAGGAGGCTGTCTCAG 323
Db 261 CCAAGTGCACCCCGGCGAGGCTCGCCCTTTGTTTGCAGTAAGAGGAGGAGGCTGTCTCAG 320
QY 324 CTGCAGAGGGGTCTATCCCTGCTTCAAGCCAGTGCCTCTTCCAGCTCCCAATGGGGACAC 383
Db 321 CTGCAGAGGGGTCTATCCCTGCTTCAAGCCAGTGCCTCTTCCAGCTCCCAATGGGGACAC 380
QY 384 CGAAGCCACGCTCCGATGAAACGTTGACGTGAAGAGGAGGATGCAGGACGAAGATCT 443
Db 381 CGAAGCCACGCTCCGATGAAACGTTGACGTGAAGAGGAGGATGCAGGACGAAGATCT 440
QY 444 TCCAGGCCACTCCAGAGAGAGCGGGGGTGGAACTGCTTGGCAGCCCGGTGAAGACAC 503
Db 441 TCCAGGCCACTCCAGAGAGAGCGGGGGTGGAACTGCTTGGCAGCCCGGTGAAGACAC 500
QY 504 ATCCTCTCTCCCAACACGCTAAATTTCAACGAGCGCATCGTAAGAGGAGGAGCGTGT 563
Db 501 ATCCTCTCTCCCAACACGCTAAATTTCAACGAGCGCATCGTAAGAGGAGGAGCGTGT 560
QY 564 GGGCCAGAGATCAATTTCTCTGATCAGAGTGAAGGGTCCCTGCTGTCCGATCACTT 623
Db 561 GGGCCAGAGATCAATTTCTCTGATCAGAGTGAAGGGTCCCTGCTGTCCGATCACTT 620
QY 624 CTCTGATACCCCTGATGACCTGATATTAACGTGATGATCATCGACACCCCGATGAGAC 683
Db 621 CTCTGATACCCCTGATGACCTGATATTAACGTGATGATCATCGACACCCCGATGAGAC 680
QY 684 CGACTCGCTGGAGTTCTCTGGGAATGGCAACGAACTGGAGTGGGAAGACGACACCCCGT 743
Db 681 CGACTCGCTGGAGTTCTCTGGGAATGGCAACGAACTGGAGTGGGAAGACGACACCCCGT 740
QY 744 GGGCCAGGCAAGAACTCCCGGGAAGCGCGGATCTATTGGGGACGGCACGACGGA 803
Db 741 GGGCCAGGCAAGAACTCCCGGGAAGCGCGGATCTATTGGGGACGGCACGACGGA 800
QY 804 GGAAGGAGCGCGCCCAACGGGCGCTGTGGCGGACAGTGATCATCGGGGAGCAGAGCA 863
Db 801 GGAAGGAGCGCGCCCAACGGGCGCTGTGGCGGACAGTGATCATCGGGGAGCAGAGCA 860
QY 864 CGGTATAGACTGCAATGATCGGCTTTACATGAAGTGGTCAACCGAGGGTACTA 923
Db 861 CGGTATAGACTGCAATGATCGGCTTTACATGAAGTGGTCAACCGAGGGTACTA 920
QY 924 CGGCGAAGGCTCAAGCCCATCATGCTTTGGAGCTGCTTCTTCCAGACAGAGCT 983
Db 921 CGGCGAAGGCTCAAGCCCATCATGCTTTGGAGCTGCTTCTTCCAGACAGAGCT 980
QY 984 CCGGACTACCACTACATCATGAGAACCTCTTCTGTAGCTCATCAGCAGCTTAGGCT 1043
Db 981 CCGGACTACCACTACATCATGAGAACCTCTTCTGTAGCTCATCAGCAGCTTAGGCT 1040
QY 1044 CTGTGTGCTGAGGACTACATGATCGGTACCTGAAAGCTGCGACCGCCCGCGGAGGAT 1103
Db 1041 CTGTGTGCTGAGGACTACATGATCGGTACCTGAAAGCTGCGACCGCCCGCGGAGGAT 1100
QY 1104 GCTGTGAATCGGCTGCTGAGAGTGTCTACAGATGATCGACCGGAGGTTCCGGAAGAA 1163
Db 1101 GCTGTGAATCGGCTGCTGAGAGTGTCTACAGATGATCGACCGGAGGTTCCGGAAGAA 1160
QY 1164 CTGTAAGTCTTTGATCATGCTCCACCCCTCGTGGTTCTTTCGAGCTGTGCTGGCCATCTC 1223
Db 1161 CTGTAAGTCTTTGATCATGCTCCACCCCTCGTGGTTCTTTCGAGCTGTGCTGGCCATCTC 1220
QY 1224 TCGCCCTTTTCATCAGCGTCAAGTTTCATCAAGATCCAGTACGTGCAAGCTTGAAGA 1283
Db 1221 TCGCCCTTTTCATCAGCGTCAAGTTTCATCAAGATCCAGTACGTGCAAGCTTGAAGA 1280
QY 1284 CTGAGGCAACTCATCCCTTATGGAAAGCAAGTCCAGATCCAGATCGCTCTGCAATACGA 1343

Db 1281 CTTGAGAGCACTCATCTCTTATGGAACACGTCAGATCCAGATCGCTCTGCAATACGA 1340
QY 1344 AGAGGAAAGACTGAAGGCCAGGAGGAGGAGCGGAGGCCCGGAGCTTTGTGCTGCC 1403
Db 1341 AGAGGAAAGACTGAAGGCCAGGAGGAGGAGCGGAGGCCCGGAGCTTTGTGCTGCC 1400
QY 1404 CAGGCTCTGAAGAGAGCGAGGTGGCACCACTGAGGAAACAGGTCCTGCTCTGCTCTCAGA 1463
Db 1401 CAGGCTCTGAAGAGAGCGAGGTGGCACCACTGAGGAAACAGGTCCTGCTCTGCTCTCAGA 1460
QY 1464 AGATCAGGAAACCAAGCATGTCTCTGAGGCGAGCTGAGGATTAACAAGGACATGGAAGAAG 1523
Db 1461 AGATCAGGAAACCAAGCATGTCTCTGAGGCGAGCTGAGGATTAACAAGGACATGGAAGAAG 1520
QY 1524 TTCCAGATGCCAGAAAACCTCTCTCAGAGCGCCACTGGCCCCAGATCTCATCTGCTCTCA 1583
Db 1521 TTCCAGATGCCAGAAAACCTCTCTCAGAGCGCCACTGGCCCCAGATCTCATCTGCTCTCA 1580
QY 1584 TCTGAGTCCCAATCTTCCAGGGTGCAGGCGCTCCCGTTCTCTCTGAAAACCCAGCATC 1643
Db 1581 TCTGAGTCCCAATCTTCCAGGGTGCAGGCGCTCCCGTTCTCTCTGAAAACCCAGCATC 1640
QY 1644 CTTTTCAGTCTGCTGAAAACCATTTGATTTTTTTTTTTTTTAAAGAGGATTTGTGCGTT 1703
Db 1641 CTTTTCAGTCTGCTGAAAACCATTTGATTTTTTTTTTTTTTAAAGAGGATTTGTGCGTT 1700
QY 1704 CCAGAAAAGGGCCAGCTCTGAGCCCTCACCTTCCACACTCAGGAACTCTCAGCCGAG 1763
Db 1701 CCAGAAAAGGGCCAGCTCTGAGCCCTCACCTTCCACACTCAGGAACTCTCAGCCGAG 1760
QY 1764 GAAGCAAGAAAGCGCAGGGGGTGGCCCGCTGGGCTCGGTGGCTCCGCTCTCTGCTCGCA 1823
Db 1761 GAAGCAAGAAAGCGCAGGGGGTGGCCCGCTGGGCTCGGTGGCTCCGCTCTCTGCTCGCA 1820
QY 1824 GCGCTGTGTGTCAGAGCTGGAATAAGATTAAGAACCTTCTCTTGTGTGTCAACCGCTC 1883
Db 1821 GCGCTGTGTGTCAGAGCTGGAATAAGATTAAGAACCTTCTCTTGTGTGTCAACCGCTC 1880
QY 1884 CAGGTTGAGGACACAGACACCCAGCCCGGCTGGGTCTGGCTCTTCTCTGCTGCC 1943
Db 1881 CAGGTTGAGGACACAGACACCCAGCCCGGCTGGGTCTGGCTCTTCTCTGCTGCC 1940
QY 1944 TTTTCCCTCCAGAAATGCGGCTCAGACCTTAGAAGCTCAACCCCTCTATGAGGCGCACGTC 2003
Db 1941 TTTTCCCTCCAGAAATGCGGCTCAGACCTTAGAAGCTCAACCCCTCTATGAGGCGCACGTC 2000
QY 2004 TGGGTTAGCTCTGACCTCCGACCTTATGTCCAAATTTCAACCCATGTTTCTCATTTG 2063
Db 2001 TGGGTTAGCTCTGACCTCCGACCTTATGTCCAAATTTCAACCCATGTTTCTCATTTG 2060
QY 2064 ACCGCGCCCTTCTCGCTCATATGACACCCAGCTCTCTTGTAGAGGATCAGAGCCATG 2123
Db 2061 ACCGCGCCCTTCTCGCTCATATGACACCCAGCTCTCTTGTAGAGGATCAGAGCCATG 2120
QY 2124 CACAAGAGAGCGCTGCGCAACCATCTTGTCTCCGATTCGAAATGCAAAATGCAACCCAGTAA 2183
Db 2121 CACAAGAGAGCGCTGCGCAACCATCTTGTCTCCGATTCGAAATGCAAAATGCAACCCAGTAA 2180
QY 2184 TCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACCGGGCAAAACCCGTCAA 2243
Db 2181 TCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACCGGGCAAAACCCGTCNA 2240
QY 2244 TACCTCCCAACCAAGATGAGATATGGAATCTCATGCTCTCCCAACCCAGGCTCAGGC 2303
Db 2241 TACCTCCCAACCAAGATGAGATATGGAATCTCATGCTCTCCCAACCCAGGCTCAGGC 2300
QY 2304 TGGGACACGCCAAACGCTGTTCCGGTGGAAACAGAGGCTCAGAACTGCTCTGAAA 2363
Db 2301 TGGGACACGCCAAACGCTGTTCCGGTGGAAACAGAGGCTCAGAACTGCTCTGAAA 2360
QY 2364 TAGGCAGACTTAGCAAGAGGAGATACAGGGTATCGGGCTTTGAGTGTTCAGAACTCA 2423
Db 2361 TAGGCAGACTTAGCAAGAGGAGATACAGGGTATCGGGCTTTGAGTGTTCAGAACTCA 2420

Qy	968	TTCCAGACAGCAGCCTCCCGACTACCACTACATCATGGAGAACTCTTCTGTGACGTCA	1027
Db	960	TTCCAGACAGCAGCCTCCCGACTACCACTACATCATGGAGAACTCTTCTGTGACGTCA	1019
Qy	1028	TCAGCAGCTTAGAGCTCTGTGCGTGGAGACTACATGATCGTGTACTTGAAACGGTGCCA	1087
Db	1020	TCAGCAGCTTAGAGCTCTGTGCGTGGAGACTACATGATCGTGTACTTGAAACGGTGCCA	1079
Qy	1088	CGCCCGCGGAGGATGCTTGGAAATCGCTCGCTGGAAGAGTGTCTACAGATGATCGAAC	1147
Db	1080	CGCCCGCGGAGGATGCTTGGAAATCGCTCGCTGGAAGAGTGTCTACAGATGATCGAAC	1139
Qy	1148	GGAGGTTGCGGAAAAAAGCTGAAGTCTTTGATCATGTGTCACCCCTCTGTGTTCAATTCGGA	1207
Db	1140	GGAGGTTGCGGAAAAAAGCTGAAGTCTTTGATCATGTGTCACCCCTCTGTGTTCAATTCGGA	1199
Qy	1208	CTGTGCTGCGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTACG	1267
Db	1200	CTGTGCTGCGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTACG	1259
Qy	1268	TGCAACAGCTTGGAAAGACCTGGAGCAACTCATCCCTATTGAAACAGTCCAGATCCACAGCT	1327
Db	1260	TGCAACAGCTTGGAAAGACCTGGAGCAACTCATCCCTATTGAAACAGTCCAGATCCACAGCT	1319
Qy	1328	CGCTCTGCGAAATACGAAGAGAAAGACTGAAGCCACGAGGAGGAGCGCGAGGCCCCAGC	1387
Db	1320	CGCTCTGCGAAATACGAAGAGAAAGACTGAAGGCCACGAGGAGGAGCGCGAGGCCCCAGC	1379
Qy	1388	CGGAGTTGTGCTGCCCGAGGTTCTGAAGAGAGCCAGAGGTGGCCACAGTGTGAAAAACAGGT	1447
Db	1380	CGGAGTTGTGCTGCCCGAGGTTCTGAAGAGAGCCAGAGGTGGCCACAGTGTGAAAAACAGGC	1439
Qy	1448	CTGCTCTGGTCTCAGAAAGATCAGGAAACAAAGCATGTCTCAGGCGAGCTGAGCATTAACAA	1507
Db	1440	CTGCTCTGGTCTCAGAAAGATCAGGAAACAAAGCATGTCTCAGGCGAGCTGAGCATTAACAA	1499
Qy	1508	AGCAGATGGAAGAAAGATTCCAGATGCGCAGAAAACTCTGTGCAGCGCCACTGGCCCCAG	1567
Db	1500	AGCAGATGGAAGAAAGATTCCAGATGCGCAGAAAACTCTGTGCAGCGCCACTGGCCCCAG	1559
Qy	1568	ATCTCATCTGCGCTCATCTGAGTGCCCAATCTTCCAGGGTGCAGGCCCTCGTTTCATC	1627
Db	1560	ATCTCATCTGCGCTCATCTGAGTGCCTCAAGTGCCTTCCAAAGGTGCAGGCCCTCGTTTCATC	1619
Qy	1628	TCTGAAACCCAGCATCTCTTTCAGCTGCTTTGAAAAATTTGATTTTTTTTTTTTAAACGAT	1687
Db	1620	TCTGAAACCCAGCATCTCTTTCAGCTGCTTTGAAAAATTTGATTTTTTTTTTTTAAACGAT	1679
Qy	1688	GCAGTATTTGTGCTTTCAGAAAAAGGCCAGCTCTGAGCCCTCAACCTTCCACTCA	1747
Db	1680	GCAGTATTTGTGCTTTCAGAAAAAGGCCAGCTCTGAGCCCTCAACCTTCCACTCA	1739
Qy	1748	CGAACTCTCAGCGAGAAAGCAAGAGCCAGGGGTGGCCCGTGGGTGGGTGGGC	1807
Db	1740	CGAACTCTCAGCGAGAAAGCAAGAGCCAGGGGTGGCCCGTGGGTGGGTGGGC	1799
Qy	1808	TCGCTCTCTGCTCGCAGCCCTCTGTGTCAGAGCTGGATACAAGATTCAGAGCCCTTCTCT	1867
Db	1800	TCGCTCTCTGCTCGCAGCCCTCTGTGTCAGAGCTGGATACAAGATTCAGAGCCCTTCTCT	1859
Qy	1868	TGCTTGTGACCCCGCTCCAGGTTGGAGCCACAGACCCCGCAGCCCGGCTGGGTCTG	1927
Db	1860	TGCTTGTGACCCCGCTCCAGGTTGGAGCCACAGACCCCGCAGCCCGGCTGGGTCTG	1919
Qy	1928	CGTCTTTTCTGTGCTTTTCCCTCCAGATGCGGCTTCAGACTAGAGCTCAACCCCCC	1987
Db	1920	CGTCTTTTCTGTGCTTTTCCCTCCAGATGCGGCTTCAGACTAGAGCTCAACCCCCC	1979
Qy	1988	TATGAGGGCCAGCTCTCGGGTGTAGTCTCTGACCTCCGACCTTATGTGTCAAAATTCACCC	2047
Db	1980	TATGAGGGCCAGCTCTCGGGTGTAGTCTCTGACCTCCGACCTTATGTGTCAAAATTCACCC	2039
Qy	2048	CATGGTTTTTCATTTTGACCCGCCCTTCTCGCTCATATGACACCCAGCTCTTTGAGA	2107

[illegible]

AUTHORS Ota,T., Isoqai,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 15950 07-FEB-2001;
Research Association for Biotechnology (JP)

FEATURES

source

1. .2894
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
872..1480
/note="unnamed protein product"

CDS

/codon_start=1
/protein_id="CAE91182.1"
/db_xref="GI:40035782"
/translation="MLRPMKVTHGGYEGEGLNLIIVPAACFLPDSSLPDYHYIMEN
LFLVTSLSLELLVAEDYMLVNLGATPRRMPGICWKKCYQVIDRLRLKXLSLIIV
HESWFIPTVLAISRPISSVSKINKIQVHSLDLEQLIPEWHVQIPDCVLQYEEERLKL
ARRSARPOPEFLVPRSEBKPEVAPENRPAALVSDQETSMS"

ORIGIN

Query Match 98.4%; Score 2870.2; DB 6; Length 2894;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2886; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 8 CTCCTGCAGCCCTGAGCTGGGAAGAAGACGACTACCTCGGAGGCGAGGCGCGAGCGGCG 67
DB 1 CTCCTGCAGCCCGGGCTGGGAAGAAGACGACTACCTCGGAGGCGAGGCGCGAGCGGCG 60
QY 68 GGCAGTGAAGGGGGCGAGCGCGAGCCCGCGCTGGGAGGCCACCGCTAACCCCTGCAC 127
DB 61 GGCAGTGAAGGGGGCGAGCGCGAGCCCGCGCTGGGAGGCCACCGCTAACCCCTGCAC 120
QY 128 CCCACCCACCCCTGCACAAAGAGCTGGGGGCGTGGCCACGCTCGCCCTGGGTGACCTT 187
DB 121 CCCACCCACCCCTGCACAAAGAGCTGGGGGCGTGGCCACGCTCGCCCTGGGTGACCTT 180
QY 188 CCTCGGATGAGAAATCGCCCTCGGAGCATCTCTTCCTCTAGGCTCTGAAGGCGCGG 247
DB 181 CCTCGGATGAGAAATCGCCCTCGGAGCATCTCTTCCTCTAGGCTCTGAAGGCGCGG 240
QY 248 GGCAGTGAAGGGGGCGAGCGCGAGCTGCAACCGGGGAGGCTGCTTTGTTGCGCATGAAGA 307
DB 241 GGCAGTGAAGGGGGCGAGCGCGAGCTGCAACCGGGGAGGCTGCTTTGTTGCGCATGAAGA 300
QY 308 GGCAGGCTGCTCTCAGCTGAGAGGGGTCACTCCCTCTCAAGCCAGTGCCTCTCCAG 367
DB 301 GGCAGGCTGCTCTCAGCTGAGAGGGGTCACTCCCTCTCAAGCCAGTGCCTCTCCAG 360
QY 368 CTCCCATGGGACCAACGAGCCACGCTCGGATGGAAACGTTGGACGTTGAAGGAGGAAAT 427
DB 361 CTCCCATGGGACCAACGAGCCACGCTCGGATGGAAACGTTGGACGTTGAAGGAGGAAAT 420
QY 428 GGCAGGACGAGATCTTCCAGGCGACTCCAGAGAGAGCGGGGTGGAACTGCTTTGGCA 487
DB 421 GGCAGGACGAGATCTTCCAGGCGACTCCAGAGAGAGCGGGGTGGAACTGCTTTGGCA 480
QY 488 GCCCGGTGGAGACACATCTCTCTCCCAACAGCTTAAATTTCAACGGAGCGCATCTTA 547
DB 481 GCCCGGTGGAGACACATCTCTCTCCCAACAGCTTAAATTTCAACGGAGCGCATCTTA 540
QY 548 AGAGGAAGCGCTGGTGGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGAAGGGTCCC 607
DB 541 AGAGGAAGCGCTGGTGGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGAAGGGTCCC 600
QY 608 TGCCTCCGATGACTTTTGGATACCCCTGATGACCTGGATTAATTAACGTGATGATCG 667
DB 601 TGCCTCCGATGACTTTTGGATACCCCTGATGACCTGGATTAATTAACGTGATGATCG 660
QY 668 AGACCCCGGATGAGACCGACTCGCTCGAGTTCTCTGGGGAATGGCAACGAACCTGGAGTGG 727
DB 661 AGACCCCGGATGAGACCGACTCGCTCGAGTTCTCTGGGGAATGGCAACGAACCTGGAGTGG 720

QY 728 AAGACGACACCCCGTGGCCACCGCCAAAGAACATCCCGGGGACAGCGGATCTATTG 787
DB 721 AAGACGACACCCCGTGGCCACCGCCAAAGAACATCCCGGGGACAGCGGATCTATTG 780
QY 788 GGGACGGACACGACGAGAGAGCGGCGGCGCAACGGGGCTGTGTGGGAGACAGTGATCA 847
DB 781 GGGACGGACACGACGAGAGAGCGGCGGCGCAACGGGGCTGTGTGGGAGACAGTGATCA 840
QY 848 TCGGGGAGCAGAGACACCGTATAGACCTGACATCATCGGCTTACATGAAGTGGTCA 907
DB 841 TCGGGGAGC-AGAGCACCGTATAGACCTGACATCATCGGCTTACATGAAGTGGTCA 899
QY 908 CCCACGAGGGTACTACGCGGAAGGCTCAAGCCCATCATCGTCTTCGAGCGCTGCTCC 967
DB 900 CCCACGAGGGTACTACGCGGAAGGCTCAAGCCCATCATCGTCTTCGAGCGCTGCTCC 959
QY 968 TTCCAGACAGCAGCCCTCCCGACTACCATCATCATGAGAGAACCTCTTCTGTGATGCA 1027
DB 960 TTCCAGACAGCAGCCCTCCCGACTACCATCATCATGAGAGAACCTCTTCTGTGATGCA 1019
QY 1028 TCAGCAGCTTACAGCTCTCTGTGGCTGAGGACTACATCATCGTGTACCTGAACGGTGC 1087
DB 1020 TCAGCAGCTTACAGCTCTCTGTGGCTGAGGACTACATCATCGTGTACCTGAACGGTGC 1079
QY 1088 CGCCCCGGCGGAGGATGCTCGGAATCGGCTGGCTGAAGAAGTGTCTACAGATGATCGAC 1147
DB 1080 CGCCCCGGCGGAGGATGCTCGGAATCGGCTGGCTGAAGAAGTGTCTACAGATGATCGAC 1139
QY 1148 GGAGGTTGGGAAACCTCGAAGTCTTGTATCATCGTCCACCTCTGTGGTTCATTGGA 1207
DB 1140 GGAGGTTGGGAAACCTCGAAGTCTTGTATCATCGTCCACCTCTGTGGTTCATTGGA 1199
QY 1208 CTGTCTCGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCATCAACAGATCCAGTAG 1267
DB 1200 CTGTCTCGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCATCAACAGATCCAGTAG 1259
QY 1268 TGACAGCTTGGAGACCTGGAGCAATCATCTCTTATGAAACACCTCCAGATCCAGACT 1327
DB 1260 TGACAGCTTGGAGACCTGGAGCAATCATCTCTTATGAAACACCTCCAGATCCAGACT 1319
QY 1328 GGGTCTCTGAATACGAAGAGGAAGATGTAAGGCCAGGAGGAGAGCGGAGGCCCGAGC 1387
DB 1320 GGGTCTCTGAATACGAAGAGGAAGATGTAAGGCCAGGAGGAGAGCGGAGGCCCGAGC 1379
QY 1388 CGGAGTTGTGTCTGCCAGGCTCTGAAGAGAGCCAGAGGTGGCACAGTGGAAAAACAGT 1447
DB 1380 CGGAGTTGTGTCTGCCAGGCTCTGAAGAGAGCCAGAGGTGGCACAGTGGAAAAACAGC 1439
QY 1448 CTGCTCTGCTCAGAGATCAGGAAACAGCATGTCTCTGAGGGGACGTTGAGCATAACAA 1507
DB 1440 CTGCTCTGCTCAGAGATCAGGAAACAGCATGTCTCTGAGGGGACGTTGAGCATAACAA 1499
QY 1508 AGGACATGGAAGAGATTCAGATGCCAGAAAACTCTGTCTGACGCGCCCATCTGGCCCC 1567
DB 1500 AGGACATGGAAGAGATTCAGATGCCAGAAAACTCTGTCTGACGCGCCCATCTGGCCCC 1559
QY 1568 ATCTCATCTCTGCTCATCTGAGTCCCAATCTTCCAGGGTGCAGCGCCCTCGTTCATC 1627
DB 1560 ATCTCATCTCTGCTCATCTGAGTCCCAATCTTCCAGGGTGCAGCGCCCTCGTTCATC 1619
QY 1628 TCTGAAACCCAGCATCTCTTTCAGCTGTCTGAAAAATTTGTTATTTTTTTTTTAAACGAT 1687
DB 1620 TCTGAAACCCAGCATCTCTTTCAGCTGTCTGAAAAATTTGTTATTTTTTTTTTAAACGAT 1679
QY 1688 GCAGTATTTGTGCTTCCAGAAAAAGGGGCGCAGCTCTGAGCCCTCACCTTCCACACTCA 1747
DB 1680 GCAGTATTTGTGCTTCCAGAAAAAGGGGCGCAGCTCTGAGCCCTCACCTTCCACACTCA 1739
QY 1748 CGAACTCTCAGCCGAGGAGGCAAGCAAGCGAGGGGGTGGCCCCGCTGCGTGGTGGCC 1807
DB 1740 CGAACTCTCAGCCGAGGAGGCAAGCAAGCGAGGGGGTGGCCCCGCTGCGTGGTGGCC 1799
QY 1808 TCCGCTCTGCTCGCAGGCCCTGTGTGTCAGAGCTGGATACAAGATTCAAGACCTTCTCT 1867

Db 1800 TCGCTCTGCTCGACCCCTGTGGTCAGAGCTGGATACAGATTCAGACCTCTCTCT 1859
QY 1868 TGCTGTGTACCGCTCCAGGTGTGAGCCACAGACACCCACCGCCACCCCGCTGGGTCTG 1927
Db 1860 TGCTGTGTACCGCTCCAGGTGTGAGCCACAGACACCCACCGCCACCCCGCTGGGTCTG 1919
QY 1928 CGTCTCTTCTGTGCTCTTCTCCAGAAATGCGGCTCTCAGACTAGAGCTCAACCCCCC 1987
Db 1920 CGTCTCTTCTGTGCTCTTCTCCAGAAATGCGGCTCTCAGACTAGAGCTCAACCCCCC 1979
QY 1988 TATGAGGCCACGCTCTGCGGTAGCTCTCAGCTCCGACCTTATGTCCTCAATTTTCACAC 2047
Db 1980 TATGAGGCCACGCTCTGCGGTAGCTCTCAGCTCCGACCTTATGTCCTCAATTTTCACAC 2039
QY 2048 CATGTTTCTTCTTATGACCGCCCTCTCTCGCTCATATGACACCCAGCTCTCTTTGAGA 2107
Db 2040 CATGTTTCTTCTTATGACCGCCCTCTCTCGCTCATATGACACCCAGCTCTCTTTGAGA 2099
QY 2108 GGATCAGAGCCCATTTGCAAGAGAGCGCTGCAACCATCTTGTCTCGATTCGAA 2167
Db 2100 GGATCAGAGCCCATTTGCAAGAGAGCGCTGCAACCATCTTGTCTCGATTCGAA 2159
QY 2168 AATGACACCCAGCTTATGACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2227
Db 2160 AATGACACCCAGCTTATGACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2219
QY 2228 GGGCAAAACCCCGTCAATACCTCCACCAAGGAATGATGATGGACCTCACTGTCCTCC 2287
Db 2220 GGGCAAAACCCCGTCAATACCTCCACCAAGGAATGATGATGGACCTCACTGTCCTCC 2279
QY 2288 CAACCCAGCGTCAAGCTGGGACACGCAACGCTGTTCCGGTGTGGAACAGCAGGCTCA 2347
Db 2280 CAACCCAGCGTCAAGCTGGGACACGCAACGCTGTTCCGGTGTGGAACAGCAGGCTCA 2339
QY 2348 GAACTGGCTCTCAATAGGACACCTAGCAGAGGAAGATCAGGGTATCGGGCGTTTG 2407
Db 2340 GAACTGGCTCTCAATAGGACACCTAGCAGAGGAAGATCAGGGTATCGGGCGTTTG 2399
QY 2408 AGTGTCTTCAAGCTCAATGGGAAGATTAATCAGTGGCTGCGCCGACGCTCACTGCT 2467
Db 2400 AGTGTCTTCAAGCTCAATGGGAAGATTAATCAGTGGCTGCGCCGACGCTCACTGCT 2459
QY 2468 CAAAGCTTGGACAGCGGGTCTTGTGTCGGAGGCAATTTCCCTAGGAAAAGAGACA 2527
Db 2460 CAAAGCTTGGACAGCGGGTCTTGTGTCGGAGGCAATTTCCCTAGGAAAAGAGACA 2519
QY 2528 GACTTTTCTAATGTGTCMAATGCGGATCACTGGTCAGATGACTCTAGAGCACTGAG 2587
Db 2520 GACTTTTCTAATGTGTCMAATGCGGATCACTGGTCAGATGACTCTAGAGCACTGAG 2579
QY 2588 CTCCTGTCTCTGGAAGTATTAAGAAAAGGCTGGCCAGGACGATGCTCAAGCTGT 2647
Db 2580 CTCCTGTCTCTGGAAGTATTAAGAAAAGGCTGGCCAGGACGATGCTCAAGCTGT 2639
QY 2648 AATCCAGACTTTGGAGGCGCAGGAGCGGATCACTGAGTGAAGGATTTGAGAACA 2707
Db 2640 AATCCAGACTTTGGAGGCGCAGGAGCGGATCACTGAGTGAAGGATTTGAGAACA 2699
QY 2708 GCTGCGCAACATGTTGAACTCTATCTTCTAATAAATAAATAATAGCCAGCGGTG 2767
Db 2700 GCTGCGCAACATGTTGAACTCTATCTTCTAATAAATAAATAATAGCCAGCGGTG 2759
QY 2768 TGSCAGTCTCTGTAATCCNGCTACTTGGAGGCTGAGGATGAGATCACTTAACT 2827
Db 2760 TGSCAGTCTCTGTAATCCNGCTACTTGGAGGCTGAGGATGAGATCACTTAACT 2819
QY 2828 GAGAGCAGAGGTTTACAGTGAGCCAGATCTGTCACCTGCTTCCAGCTGCGGACAG 2887
Db 2820 GAGAGCAGAGGTTTACAGTGAGCCAGATCTGTCACCTGCTTCCAGCTGCGGACAG 2879
QY 2888 GCAAGACTCTGTCT 2902
|||||

Db 2880 GCAAGACTCTGTCT 2894

RESULT 8

AK027889

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AK027889 Homo sapiens cDNA FLJ14983 fis, clone Y79AA1000268, weakly similar to Mus musculus Nip21 mRNA.

AK027889 2894 bp mRNA linear PRI 30-JAN-2004
AK027889.1 GI:14042891
oligo capping; fis (full insert sequence).

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahori, K., Murakami, K., Tasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Kanda, Y., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Iehibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotsuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikemura, Y., Okamoto, S., Okitani, R., Kawakami, K.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, T., Shirai, Y., Wakahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.

Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

2 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,
Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,
Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,
Ishii, S., Kawai, T., Saito, K., Yamamoto, J., Wakamatsu, A.,
Nakamura, Y., Nagahori, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.

Unpublished
3 (bases 1 to 2894)
Isogai, T. and Otsuki, T.

Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

TITLE

COMMENT

FEATURES		Location/Qualifiers	
source		1..2894	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="Y79AA1000268"	
		/cell_line="Y79"	
		/cell_type="retinoblastoma"	
		/clone_lib="Y79AA1"	
		/note="cloning vector: pME18SFL3"	
ORIGIN			
Query Match 98.4%; Score 2870.2; DB 8; Length 2894;			
Best Local Similarity 99.7%; Pred. No. 0;			
Matches 2886; Conservative 0; Mismatches 8; Indels 1; Gaps 1;			
QY	8	CTCTGCCAGCCCTGAGCTGGGAAGAAGCAGCTACCTCGGAGGCGAGGCGCGGCGGCG 67	
DB	1	CTCTGCCAGCCCGGGCTGGGAAGAAGCAGCTACCTCGGAGGCGAGGCGCGGCGGCG 60	
QY	68	GGCGATGAGAGGGGGCGCAGCGCAGCCCGCGCTGGGAGGCCACCGCTAACCCCTGCAC 127	
DB	61	GGCGATGAGAGGGGGCGCAGCGCAGCCCGCGCTGGGAGGCCACCGCTAACCCCTGCAC 120	
QY	128	CCCAACCAACCCCTGCACAAAGAGCTGGGGGGCTGGGCCACGTCGCCCTGGGTGACCTT 187	
DB	121	CCCAACCAACCCCTGCACAAAGAGCTGGGGGGCTGGGCCACGTCGCCCTGGGTGACCTT 180	
QY	188	CCTCGGATGAGAAATCCGCCCTCGGAGGATCCTCTTCTCTTAGGCTCTGAGGCGCGG 247	
DB	181	CCTCGGATGAGAAATCCGCCCTCGGAGGATCCTCTTCTCTTAGGCTCTGAGGCGCGG 240	
QY	248	GGAGCGTGACGATGCCAGCTGCACCCGGGCGAGGCTCGCCCTTGTGTTGCCAGTAAGGA 307	
DB	241	GGAGCGTGACGATGCCAGCTGCACCCGGGCGAGGCTCGCCCTTGTGTTGCCAGTAAGGA 300	
QY	308	GGAGAGGCTGTCTCAGCTGCAGAGGGGTATCCCTGCTTCAAGCCAGTGCCTCTTCCCGAG 367	
DB	301	GGAGAGGCTGTCTCGGCTGCAGAGGGGTATCCCTGCTTCAAGCCAGTGCCTCTTCCCGAG 360	
QY	368	CTCCCATGGGACCGGAGCCAGCTCCGGATGGAAAACTGGACGTGAAGGAGGAAT 427	
DB	361	CTCCCATGGGACCGGAGCCAGCTCCGGATGGAAAACTGGACGTGAAGGAGGAAT 420	
QY	428	GGCAGGACGAAGATCTTCCAGGCGCACTCCAGAGAGACGGGGGTGGAACTGCTTGGCA 487	
DB	421	GGCAGGACGAAGATCTTCCAGGCGCACTCCAGAGAGACGGGGGTGGAACTGCTTGGCA 480	
QY	488	GCCCGGTGGAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCGCATCGTA 547	
DB	481	GCCCGGTGGAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCGCATCGTA 540	
QY	548	AGAGGAAGACGCTGGTGGGCCCGCAGATCAACATTTCTTGGATCAGAGTGGGGTCCC 607	
DB	541	AGAGGAAGACGCTGGTGGGCCCGCAGATCAACATTTCTCTGGATCAGAGTGGGGTCCC 600	
QY	608	TGCTGTCCGATGACTTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG 667	
DB	601	TGCTGTCCGATGACTTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACATCG 660	
QY	668	AGACCCCGATGAGACCGACTCGCTGGAGTTCTCGGGGAATGGCAACGAACTGGAGTGGG 727	
DB	661	AGACCCCGATGAGACCGACTCGCTGGAGTTCTCGGGGAATGGCAACGAACTGGAGTGGG 720	
QY	728	AAGACGACACCCCGTGGCCACCGCCAGAAACATGCCCGGGGACAGCGCGGATCTATTGG 787	
DB	721	AAGACGACACCCCGTGGCCACCGCCAGAAACATGCCCGGGGACAGCGCGGATCTATTGG 780	
QY	788	GGGACGCGACGACGAGGAGCGGACGCGCGCCCAACGGGCGCTGTGGCGGACAGTGATCA 847	
DB	781	GGGACGCGACGACGAGGAGCGGACGCGCGCCCAACGGGCGCTGTGGCGGACAGTGATCA 840	
QY	848	TCGGGGAGCAGAGCACCGGTATAGACTGCAATGATCCGGCCCTTACATGAAGTGGTCA 907	
DB			
DB	841	TCGGGGAGC-AGAGCACCGTATAGACCTGCAATGATCGGCTTACATGAAGTGGTCA 899	
QY	908	CCACGGAGGGTACTACCGCGAAGGCCCTCAAGCCCATCATCGTCTTCGACAGCTGCTTCC 967	
DB	900	CCACGGAGGGTACTACCGCGAAGGCCCTCAAGCCCATCATCGTCTTCGACAGCTGCTTCC 959	
QY	968	TTCCAGACAGCAGCCTCCCGCACTACCATCATCATGAGAGAACCTCTTCTGTGATGCA 1027	
DB	960	TTCCAGACAGCAGCCTCCCGCACTACCATCATCATGAGAGAACCTCTTCTGTGATGCA 1019	
QY	1028	TCAGCAGCTTAGAGCTCCTGTGGTGGAGTACTACATGATCGTGTACTGAACCGTGGCA 1087	
DB	1020	TCAGCAGCTTAGAGCTCCTGTGGTGGAGTACTACATGATCGTGTACTGAACCGTGGCA 1079	
QY	1088	CGCCCGCGGAGGATGCTGGAAATCGGCTGGCTGAAGAGTGTCTACAGATGATCGACC 1147	
DB	1080	CGCCCGCGGAGGATGCTGGAAATCGGCTGGCTGAAGAGTGTCTACAGATGATCGACC 1139	
QY	1148	GGAGGTTTCGGAAAAACCTGAAATCCTTGAATCATCGTCCACCCCTCGTGGTTCATTTCGA 1207	
DB	1140	GGAGGTTTCGGAAAAACCTGAAATCCTTGAATCATCGTCCACCCCTCGTGGTTCATTTCGA 1199	
QY	1208	CTGTCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTAGC 1267	
DB	1200	CTGTCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTAGC 1259	
QY	1268	TGCACAGCTTGGAGACCTGGAGCAATCATCCCTATCGAAACACCTGATCCAGATCCCAGACT 1327	
DB	1260	TGCACAGCTTGGAGACCTGGAGCAATCATCCCTATCGAAACACCTGATCCAGATCCCAGACT 1319	
QY	1328	GGGTCTCTGCAATACCAAGAGGAAAGACTGAAGGCCAGAGGGAGAGCGCGAGGCCCCAGC 1387	
DB	1320	GGGTCTCTGCAATACCAAGAGGAAAGACTGAAGGCCAGAGGGAGAGCGCGAGGCCCCAGC 1379	
QY	1388	CGGAGTGTGTGTCGCCAGGTCCTGAAGAGAGCCAGAGGTGGCACCACTGTGGAATAACAGGT 1447	
DB	1380	CGGAGTGTGTGTCGCCAGGTCCTGAAGAGAGCCAGAGGTGGCACCACTGTGGAATAACAGC 1439	
QY	1448	CTGCTCTGGTCTCAGAAATCAGGAAACAAGCATGTCTCTGAGGCGACGTGAGCATAACAA 1507	
DB	1440	CTGCTCTGGTCTCAGAAATCAGGAAACAAGCATGTCTCTGAGGCGACGTGAGCATAACAA 1499	
QY	1508	AGGACATGGAAGAAGATTCCAGATGCCAGAAAACTCTGTGACAGCGCCACTTGGCCCCAG 1567	
DB	1500	AGGACATGGAAGAAGATTCCAGATGCCAGAAAACTCTGTGACAGCGCCACTTGGCCCCAG 1559	
QY	1568	ATCTCATCTGCTCATCTGAGTCCCCTTTCCTCAAGGGTGCAGCCCTCCGTTTCATC 1627	
DB	1560	ATCTCATCTGCTCATCTGAGTCCCCTTTCCTCAAGGGTGCAGCCCTCCGTTTCATC 1619	
QY	1628	TCTGAAACCCAGCATCTTTTCAGCTGCTTGAACAACTTGATTTTTTTTTTTTAAAGAT 1687	
DB	1620	TCTGAAACCCAGCATCTTTTCAGCTGCTTGAACAACTTGATTTTTTTTTTTTAAAGAT 1679	
QY	1688	GCAGTATTGTGCGTTCCAGAAAAAGGCGCCAGCTCTGAGCCCTTCACTTCCACACTCA 1747	
DB	1680	GCAGTATTGTGCGTTCCAGAAAAAGGCGCCAGCTCTGAGCCCTTCACTTCCACACTCA 1739	
QY	1748	CGAATCTCAGCCGAGGAGGCAAGAACGCGAGGGGTGGCCCGCTGGCTGGGTGGCC 1807	
DB	1740	CGAATCTCAGCCGAGGAGGCAAGAACGCGAGGGGTGGCCCGCTGGCTGGGTGGCC 1799	
QY	1808	TCGGCTCTGCTCGCAGCCCTTGTGGTTCAGAGCTGGATACAGATTCAAGACCTTCTCT 1867	
DB	1800	TCGGCTCTGCTCGCAGCCCTTGTGGTTCAGAGCTGGATACAGATTCAAGACCTTCTCT 1859	
QY	1868	TGCTTGTCAACCGCTTCCAGGTTGGAGCCACAGACACCAACCGCCACCCCGCTGGGTCTG 1927	
DB	1860	TGCTTGTCAACCGCTTCCAGGTTGGAGCCACAGACACCAACCGCCACCCCGCTGGGTCTG 1919	
QY	1928	CGTCTCTTCTGTGCTTTCCTCCAGAAATGCGGCTCAGACCTTAGAAGCTCAACCCCCC 1987	

Db 1920 CGTCCTTCTGTCCTTCCCTCCAGAAATGGCGCTCAGACCTAGAGCTCAACCCCC 1979
QY 1988 TATGAGGCGCCACTCTCTGGGGTAGCTCTCGACCTCGACCTATGTCCTCAAAATTTACACC 2047
Db 1980 TATGAGGCGCCACTCTCTGGGGTAGCTCTCGACCTCGACCTATGTCCTCAAAATTTACACC 2039
QY 2048 CATGGTTTTCATTTGACCGCCCTTCTCGCTCATATGACACCGAGTCTCTTTGAGA 2107
Db 2040 CATGGTTTTCATTTGACCGCCCTTCTCGCTCATATGACACCGAGTCTCTTTGAGA 2099
QY 2108 GATCAGAGCCCATTTGACAGAGAGCGCTGCTCAACCATCTCTGCTCTCGATGCA 2167
Db 2100 GATCAGAGCCCATTTGACAGAGAGCGCTGCTCAACCATCTCTGCTCTCGATGCA 2159
QY 2168 ATGACACCCAGTAATCTAGAACTTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2227
Db 2160 ATGACACCCAGTAATCTAGAACTTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2219
QY 2228 GGGCAACCCCGTCAATACCTCCCAAGGAATGAGATATGAGACCTCACTGTCTCCC 2287
Db 2220 GGGCAACCCCGTCAATACCTCCCAAGGAATGAGATATGAGACCTCACTGTCTCCC 2279
QY 2288 CAACCCAGGCTCAGGCTGGGACACGCAACGCTGTTCCGGTTGGACAGCAGGCTCA 2347
Db 2280 CAACCCAGGCTCAGGCTGGGACACGCAACGCTGTTCCGGTTGGACAGCAGGCTCA 2339
QY 2348 GAACTGGCTCTGAATAGGCAAGCTTAGCAAGAGGAATACAGGGTATCGGGTTTG 2407
Db 2340 GAACTGGCTCTGAATAGGCAAGCTTAGCAAGAGGAATACAGGGTATCGGGTTTG 2399
QY 2408 AGTGTTCAGAGTCAATTCGGGAAGATAATCCAGTCCGCTGGCCGACCACTGCAAT 2467
Db 2400 AGTGTTCAGAGTCAATTCGGGAAGATAATCCAGTCCGCTGGCCGACCACTGCAAT 2459
QY 2468 CAAGCTTCGACAGCGGCTTCTGTTCCGGAGGCAATTTCCCTAGGAAAAGAGACA 2527
Db 2460 CAAGCTTCGACAGCGGCTTCTGTTCCGGAGGCAATTTCCCTAGGAAAAGAGACA 2519
QY 2528 GACTTTCTTAATGTGTCCAAATGCGGATCACTGGTTCAGATGACCTTAGAAGCACTGAG 2587
Db 2520 GACTTTCTTAATGTGTCCAAATGCGGATCACTGGTTCAGATGACCTTAGAAGCACTGAG 2579
QY 2588 CTCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACGATGCTCACGCTGT 2647
Db 2580 CTCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACGATGCTCACGCTGT 2639
QY 2648 AATCCAGACTTTGGGAGCGGAGGAGCGGATCACCTGAGGTGAGGATTTGAGAAC 2707
Db 2640 AATCCAGACTTTGGGAGGTGAGGAGCGGATCACCTGAGGTGAGGATTTGAGAAC 2699
QY 2708 GCCTGGCCAAATGGTGAACCTCTATCTCTAATAAATACAAAATAGCCAGCGTGG 2767
Db 2700 GCCTGGCCAAATGGTGAACCTCTATCTCTAATAAATACAAAATAGCCAGCGTGG 2759
QY 2768 TGGCAGGTGCTGTATATCCAGCTACTTGGGAGGCTGAGGATGAGAACTCACTAAACCT 2827
Db 2760 TGGCAGGTGCTGTATATCCAGCTACTTGGGAGGCTGAGGATGAGAACTCACTAAACCT 2819
QY 2828 GAGAGCGAGGTTTACAGTACGAGCAAGATCGTGCCACTGCAATTCAGCTGGGCGACAGA 2887
Db 2820 GAGAGCGAGGTTTACAGTACGAGCAAGATCGTGCCACTGCAATTCAGCTGGGCGACAGA 2879
QY 2888 GCAAGACTCTGTCTC 2902
Db 2880 GCAAGACTCTGTCTC 2894

RESULT 9
AX747472
LOCUS AX747472 2759 bp mRNA linear PAT 20-JUN-2003
DEFINITION Sequence 997 from Patent EP1308459.
ACCESSION AX747472
VERSION AX747472.1 GI:32131860

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE

AUTHORS

Isoqai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and
Masuno, Y.

Full-length cDNA sequences

Patent: EP 1308459-A 997 07-MAY-2003;

Helix Research Institute (JP) ; Research Association for

Biotechnology (JP)

FEATURES

source

1..2759

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 89.0%; Score 2598.2; DB 6; Length 2759;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2756; Conservative 0; Mismatches 3; Indels 136; Gaps 2;

QY 8 CTCTGCCAGCCCTGAGCTGGGAAGAGCAGTACCTCGAGGCGCGGCGCGGCG 67
Db 1 CTCTGCCAGCCCTGAGCTGGGAAGAGCAGTACCTCGAGGCGCGGCGCGGCG 60

QY 68 GGCAGTATGAGAGGGGCGCGAGCGCGCGCGCTGGGGGCGCCACCGCTAACCTGCAC 127
Db 61 GGCAGTATGAGAGGGGCGCGAGCGCGCGCGCTGGGGGCGCCACCGCTAACCTGCAC 120

QY 128 CCCACCCACCCCTGTCACAAAAGAGCTGGCGGCGCTGGCCACGTCGCCCTGGGTGACCTT 187
Db 121 CCCACCCCA----- 128

QY 188 CCTCGGATGAGAAATCCGCCCTGGGAGGATCTCTTCTCTAGGCTCTGAAGCCCGG 247
Db 129 -----CCCTCGGAGCATCTCTTCTCTAGGCTCTGAAGCCCGG 170

QY 248 GGAGCGTGCAGCGATGCCAGCTGCACCGCGGCGAGGCTCGGCTTGTGTCAGTAAGA 307
Db 171 GGAGCGTGCAGCGATGCCAGCTGCACCGCGGCGAGGCTCGGCTTGTGTCAGTAAGA 230

QY 308 GGAGCGTGTCTCAGCTGCAGAGGGGTTCATCCCTGCTTCAAGCCAGTCCCTTCCCGAG 367
Db 231 GGAGCGTGTCTCAGCTGCAGAGGGGTTCATCCCTGCTTCAAGCCAGTCCCTTCCCGAG 290

QY 368 CTCCTATGGGACCAACCGAAGCCAGCTCCCGATGGAAAAAGCTGGAAGGAGGAGAT 427
Db 291 CTCCTATGGGACCAACCGAAGCCAGCTCCCGATGGAAAAAGCTGGAAGGAGGAGAT 350

QY 428 GGCAGGACGAGATCTTCCAGGCCCATCTCCAGAGAGACGCGGGGTGGAACTGCTTGGCA 487
Db 351 GGC----- 353

QY 488 GCCCGGTGGAAGACACATCTCTCTCCCAACACGCTAAATTTCAACGAGCGCATCGTA 547
Db 354 -----AGACACATCTCTCTCTCCCAACACGCTAAATTTCAACGAGCGCATCGTA 404

QY 548 AGAGGAAGACGCTGTGGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGAAGGAGTCC 607
Db 405 AGAGGAAGACGCTGTGGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGAAGGAGTCC 464

QY 608 TGTCTGCGATGACTTCTTGATACCCCTGATGATACCTGGATATTAACTGGATGACATCG 667
Db 465 TGTCTGCGATGACTTCTTGATACCCCTGATGATACCTGGATATTAACTGGATGACATCG 524

QY 668 AGACCCCGGATGAGACCGAGCTCGCTGGAGTTCTCTGGGGGAATGGCAACGACTGGAGTGG 727
Db 525 AGACCCCGGATGAGACCGAGCTCGCTGGAGTTCTCTGGGGGAATGGCAACGACTGGAGTGG 584

Qy	2828	GAGGGCAGAGGTTACAGTCAGCCAGATCGTGCCACTGCACTTCAGCCTGGGCGCAGAGA	2887
Db	2685	GAGGGCAGAGGTTACAGTCAGTCAGCCAGATCGTGCCACTGCACTTCAGCCTGGGCGCAGAGA	2744
Qy	2888	GCAAGACTCTGTCTC	2902
Db	2745	GCAAGACTCTGTCTC	2759

RESULT 11

AB052149	AB052149	2787 bp	linear	PRI 01-MAR-2005
LOCUS	Macaca fascicularis brain cDNA, clone:QcCE-20783, similar to human ceytaxin (ATCAY), mRNA, NM 033064.			
DEFINITION				

ACCESSION	AB052149
VERSION	AB052149.1
KEYWORDS	fis (full insert sequence).
SOURCE	Macaca fascicularis (crab-eating macaque)
ORGANISM	Macaca fascicularis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.

1 *Leucoproctaceae*, *Leucoproctinae*, Matsuda.

REFERENCE
AUTHORS Osada, N., Hida M., Kusuda, J., Tanuma, R., Iseki, K., Hirata, M., Suto, Y., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.

TITLE Assignment of 118 novel cDNAs of cynomolgus monkey brain to human chromosomes

JOURNAL Gene 275 (1), 31-37 (2001)

PURNED	11574149	(157 51 51 15002)
REFERENCE	2 (bases 1 to 2787)	
AUTHORS	Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.	
TITLE	Direct Substitution	

JOURNAL
Submitted (05-DEC-2000) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan
(E-mail: khash@nih.go.jp, URL: <http://www.nih.go.jp/yoken/genebank/>, Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)

Tel: 81-3-3263-3121 (ex. 2120), FAX: 81-3-3263-1181)
 Lab host: TOP10
 Vector: pME18S-FL3 (Acc. No. AB009864)
 R. Site1: DraIII (CACTGTGTG)
 R. Site2: DraIII (CACCATGTG)
 Description: 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
 using specific 5' and 3' primers and amplified by PCR. The PCR
 product was digested with SfiI and size selection by PCR. The PCR
 exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
 into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
 the DraIII sites can be used to isolate the cDNA insert. Libraries
 were constructed by Sugano et al. (University of Tokyo, Institute of
 Medical Science). Custom primer used for sequencing
 (5' end primer [CTTCTGCTCTAAAGCTCGC] ;
 3' end primer [CGACCTCGAGCTCAGCAGCA]).

```

FEATURES
  source
    3' end primer [CGACCTGCAGCTTCGAGCACA] ).
    Location/Qualifiers
      1..2787
        /organism="Macaca fascicularis"
        /mol_type="mRNA"
        /db_xref="taxon:9541"
        /clone="QCC8-20783"
        /sex="male"
        /issue_type="cerebellum cortex"
        /clone_lib="macaque brain cdna library QCC8"
        /dev_stage="adult"
      380..1495

```

```

CDS
7dev_stage= adult
380..1495
/note="similar to human caytaxin (ATCAY)", mRNA, NM_033064"
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB19004.1"
/db_xref="GI:11611579"
/translation="MGTTRATIRMENVVKESQEDLPRLPETSTGVELLSQPVEDT
SSPPTNLNGAHRKRKTLVADINISLDQSEGSLLSDFLTPDDLLNDVDDIETPD
ETDLSLEFLGNELWEGDDTPTVATKNMPGSADLFGDGTDEGGAAGRLWRTVIG
EQGRHLDHMTIPMYKVTGGYGEGLNAIIVPAACFLPSSLDYHYIMENLFLYV
ISSLELLVAEDATMYVLNGATPRRPMGVLWKCYOMIDRLRKNLSLITVHPSWF

```

IRTVLAISRPFISVKFINKIQYVHSLDLEQLIPMEHVQIPDCVLQYEEERLKRRES
ARPOPEFVMPRSEKKPEVAPVENRSAPVTEDOETSMS"

ORIGIN

Query Match 84.4%; Score 2463; DB 8; Length 2787;
Best Local Similarity 93.0%; Pred. No. 0;
Matches 2669; Conservative 0; Mismatches 110; Indels 90; Gaps 5;

Qy	1	GC	CGAGCCTCTG	CAGCCCTG	AGCTG	GGGAGAA	GACGTACT	CTCGAGGAG	CAGCGCGCGCA	60
Db	8	GC	CAGCCTCTG	CTAGCCTG	AGCTG	GGGAGAA	GACGTACT	CTCGAGGAG	CAGCGCGCGCA	67
Qy	61	GG	CGGCGGG	CGATG	AGAGGGGG	CGCAGCG	CCCGCGCT	TGGGAG	CCCAACCGCTAAC	120
Db	68	GG	CGGTGG	CGATG	AGAGCGGG	CGCAGCG	CCCGCGAG	GGGGAG	CCACCGCTAAC	127
Qy	121	CT	TGCACCC	CAACCC	CTTGCA	CAAAAGAG	CTTGGCGG	CGCTGGCG	CAACCGCTAAC	127
Db	128	CT	TGCACCC	CAACCC	CTTGCA	CAAAAGAG	CTTGGCGG	CGCTGGCG	CAACCGCTAAC	127
Qy	181	TG	ACCTTCTC	CGATG	CAGAAAT	CCGCGCTG	CGAGAT	CTCTTCTCT	TAGGCTCTGAA	240
Db	188	TG	ACCTTCTC	CGATG	CAGAAAT	CCGCGCTG	CGAGAT	CTCTTCTCT	TAGGCTCTGAA	247
Qy	241	GG	CCCGGG	AGCGT	GAGCGAT	GCACAG	CTGCAC	CCCGGAG	GGGCTCGCTTGTGGCA	300
Db	248	GG	CCCGGG	AGCGT	GAGCGAT	GCACAG	CTGCAC	CCCGGAG	GGGCTCGCTTGTGGCA	307
Qy	301	GT	AAGGAG	GAGGCTGT	CTCAGCTG	CGAGGGGT	CATCCCTG	CTTCAAGC	CAGTGCCTC	360
Db	308	GT	AAGGAG	GAGGCTGT	CTCAGCTG	CGAGGGGT	CATCCCTG	CTTCAAGC	CAGTGCCTC	367
Qy	361	TT	CCGAGCT	CCCATGGG	ACCA	CGAAGC	CA	CGCTCGGAT	TGAAACGTGAGCTGAA	420
Db	368	TT	TC	CGGCTCTAT	TGGGACCA	CAGAAG	CCAC	CGCTCGGAT	TGAAACGTGAGCTGAA	427
Qy	421	GAG	GAATGG	CAGGAC	GAAGAT	CTTCCAGG	CCACTCC	CAGAAGAG	AGCGGGTGGAACTG	480
Db	428	GAG	GAATGG	CAGGAC	GAAGAT	CTTCCAGG	CCACTCC	CAGAAGAG	AGCGGGTGGAACTG	487
Qy	481	CT	TGGCAG	CCCGT	TGAAAGACA	CATCCTCT	CTCTCC	CAACAC	CGCTAAATTTCAACGAGCG	540
Db	488	CT	TGGCAG	CCCGT	TGAAAGACA	CATCCTCT	CTCTCC	CAACAC	CGCTAAATTTCAACGAGCG	547
Qy	541	CAT	CGTAAG	GAGGAAG	CGCTGT	TGGCCCC	CAGAGAT	CAACATTTCT	CTGATCAGAGTGAG	600
Db	548	CAT	CGTAAG	GAGGAAG	CGCTGT	TGGCCCC	CAGAGAT	CAACATTTCT	CTGATCAGAGTGAG	607
Qy	601	GG	GTCCCTG	CTGTCC	GAATGAC	TTCTTGGAT	ATACCCCT	TGATGAC	CTGGATATTAACGTGGAT	660
Db	608	GG	GTCCCTG	CTGTCC	GAATGAC	TTCTTGGAT	ATACCCCT	TGATGAC	CTGGATATTAACGTGGAT	667
Qy	661	GA	CATCGAG	ACCCCG	ATGAGAC	CGGATCG	CTCGTGG	AGTTCT	TGGGGAATGCGCAACGAACTG	720
Db	668	GA	CATCGAG	ACCCCG	ATGAGAC	CGGATCG	CTCGTGG	AGTTCT	TGGGGAATGCGCAACGAACTG	727
Qy	721	GAG	TGGGAAG	AGACAC	CCCGCT	TGGCCAC	CGGCCA	GAACAT	CGCCGGGAGACGCGGAT	780
Db	728	GAG	TGGGAGAT	GACAC	CCCCCGTGG	CCACCG	CGCAAGAA	CATGCC	CGGGAGACGCGGAT	787
Qy	781	CT	ATTTTGGG	AGCGCA	CGAGCGG	AGCGCG	CCCAAC	CGGGCG	CTGTGGGGGACA	840
Db	788	CT	ATTTTGGG	AGCGCA	CGAGCGG	AGCGCG	CCCAAC	CGGGCG	CTGTGGGGGACG	847
Qy	841	GT	GATCAT	TGGGGAG	CAAGACA	CGGTATAG	ACTTG	AGCATG	ATCGGCTTTACATGAAA	900
Db	848	GT	GATCAT	TGGGGAG	CAAGACA	CGGTATAG	ACTTG	AGCATG	ATCGGCTTTACATGAAA	907
Qy	901	GT	GGTCA	CCCA	CGGAGGGT	TACTAC	CGCGGA	AGGCTCT	CAAGCCATCATCGTCTTCGACGC	960
Db	908	GT	GGTCA	CCCA	CGGAGGGT	TACTAC	CGCGGA	AGGCTCT	CAAGCCATCATCGTCTTCGACGC	967
Qy	961	TG	CTCTCT	TTCC	AGACAG	CAGAGC	CTCCCCG	ACTAC	CTACATCATGGAAGAACTCTTCTG	1020

[illegible]

Db	2044	TCACACCGGTGGTGTCTTTTCATTTGACCGCGCCCTTCTCGCTCATATAATGACACCCAGCTCC	2100
Qy	2101	TTTGAGAGGATCAGAGCCCATTCGACAGAGAGCGCTGCCAACCATCTCTTGTCTCTCCG	2160
Db	2104	TTTGAGAGGATCAGAGCCCATTCGACAGAGAGCTGCTGCCAACCATCTCTTGTCTCTCTG	2163
Qy	2161	ATTGCAAAATGACACACCCAGTAACTCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCA	2220
Db	2164	ATTGCAAAACGACACCCCGGTAACTCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTCA	2223
Qy	2221	AGGCACCGGCAAAACCCCGTCAATACTCCCAACAGGAATGAGATATGTGGACCTCCT	2280
Db	2224	AGGCACAGGCAAAACCCAGTCAATACTCCCAACAGGAATGAGATATGTGGACCTCCT	2283
Qy	2281	GCTCCCCACACCCAGGTGAGGCTGGGACACGCCAACGCTGTCGCGGTGGNAACGACG	2340
Db	2284	GCTCCCCACACCCCGGCATCAGGCTGGGACACGCCAACGCTGTCGCGGTGGNAACGACG	2343
Qy	2341	AGGCTCAGAAATCGGCTCTGAATAGGCGACCTAGCAAGAGGAAGATACAGGATATCG	2400
Db	2344	AGGCTCAGAGCTGGCTCTGAATAT---AGACCTAGCAAAAGGAGATACAGGATAT---	2396
Qy	2401	CGGTTTGAGTGTTCAGAAAGTCAATTCGGGAAGATAAATCCAGTGCCTGCGCCGACCCAC	2460
Db	2397	-----	2396
Qy	2461	CTGCATTCAAAGCTTGGACACGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAAA	2520
Db	2397	-----CCAGCGGGTCTTGTTCGGGAGGCAAAATTTCCCTTGGAAAAA	2438
Qy	2521	GAAGACAGACTTTCTAAATGTGTCCAAATGCGGATCACTGGTCAGATGCACTCTAGAAG	2580
Db	2439	GAAGACAGACTTTCTAAATGTGTCCAAATGCGGTCATCTGGTCAGATGCACTCTAGAAG	2498
Qy	2581	CACGTAGCTCCCTGTCTCTGGAAGTATTTAAAGAAAAAGGTGGGCCACGACGATGCTCA	2640
Db	2499	CAGTGAGCTCCCTGTCTCTGGAAGCATTTAAAGAAAAAGGTAGGCCGCGGACGATGACTCA	2558
Qy	2641	CGCTGTAAATCCAG-ACITTTGGAGCGCGGACGCGGATCACTCTGAGTGGAGGTT	2699
Db	2559	CGCTGTAAATCTTGGGACATTTGGGAGCGCGGAGTGGCGGATCCCTCTGAAGTCAAGGTT	2618
Qy	2700	TGAGAACAGCGCTCGGCCAACATGTGTGAACCTCATCTCTACTAAATAACAAAATTAGCC	2759
Db	2619	CGAGACCGCTCGGCCAACATGTGTGAACCTCATCTCTACTAAATAACAAAATTAGCC	2678
Qy	2760	AGGCGTGTGGCAGGTGCTGTAAATCCAGCTACTTTGGGAGGCTGAGGCATGAGATCAC	2819
Db	2679	AGGCGTGTGGCAGGTGCTGTAAATCTCAGTACTTTGGGAGGCTAAGGCACGAGATCAC	2738
Qy	2820	TTAAACCTGAGAGCGAGGTTACGTGAGCCAGATCTGTCACCTGCA	2868
Db	2739	TTAAACCTGAGAGCGAGGTTGACGTGAGCCAGATCTGTCACCTGCA	2787
RESULT 12			
LOCUS	CO842483	2786 bp	linear
DEFINITION	Sequence 1130 from Patent EP1440981.		
ACCESSION	CO842483		
VERSION	CO842483.1		
KEYWORDS	GI:50894270		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.		
TITLE	Full-length human cdna		
JOURNAL	Patent: EP 1440981-A 1130 28-JUL-2004; Research Association for Biotechnology (JP)		

FEATURES		Location/Qualifiers	
source	1..2786	/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Query Match		80.8%; Score 2357.6; DB 6; Length 2786;	
Best Local Similarity		95.2%; Pred. No. 0;	
Matches 2492; Conservative		0; Mismatches 4; Indels 122; Gaps 1;	
QY	300	AGTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCTCCCTGCTTCAAGCCAGTGCCT	359
DB	1	AGTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCTCCCTGCTTCAAGCCAGTGCCT	60
QY	360	CTTCCAGCTCCCATGGGACCAACGAGCCAGCTCCGGATGGAAAAAGTGGACGTGAA	419
DB	61	CTTCCAGCTCCCATGGGACCAACGAGCCAGCTCCGGATGGAAAAAGTGGACGTGAA	120
QY	420	GGAGGATGGCAGGACGATCTTCCAGGCCACTCCAGAGAGAGACGGGGTGGAACT	479
DB	121	GGAGGATGGCAGGACGATCTTCCAGGCCACTCCAGAGAGAGACGGGGTGGAACT	180
QY	480	GCTTGGCAGCCGGTGGAGACATCTCTCTCCCAACACGCTAAATTTCAAAGGAGC	539
DB	181	GCTTGGCAGCCGGTGGAGACATCTCTCTCCCAACACGCTAAATTTCAAAGGAGC	240
QY	540	GCATCGTAAGAGGAGACGCTGGTGGCCAGAGATCAACATTTCTCTGGATCAGAGTGA	599
DB	241	GCATCGTAAGAGGAGACGCTGGTGGCCAGAGATCAACATTTCTCTGGATCAGAGTGA	300
QY	600	GGGGTCCCTGCTGCTCGATGACTTTCTTGATACCCCTGATGACCTGGATATTAACGTGA	659
DB	301	GGGGTCCCTGCTGCTCGATGACTTTCTTGATACCCCTG-----	338
QY	660	TGCATCGAGACCCCGATGAGACGACTCGCTGGAGTTCCTGGGGAATGGCAACGAACT	719
DB	339	-----	338
QY	720	GGAGTGGGAGACGACACCCCGTGGCCACCGCAAGAACATGCCCGGAGCAGCGCGA	779
DB	339	-----ATGCCCGGAGCAGCGCGA	358
QY	780	TCTATTGGGAGCGGACGACGAGGAGACGGCAGCGCGCAACGGCGCGCTGTGGCGGAC	839
DB	359	TCTATTGGGAGCGGACGACGAGGAGACGGCAGCGCGCAACGGCGCGCTGTGGCGGAC	418
QY	840	AGTGATCATCGGGAGCAGAGACGATATAGACCTGACATGATCCGGCTTTAATGAA	899
DB	419	AGTGATCATCGGGAGCAGAGACCGATATAGACCTGACATGATCCGGCTTTAATGAA	478
QY	900	AGTGGTCAACCCAGGAGGTACTACGGGAGAGGCTCAACGCCATCATCGTCTTCGACG	959
DB	479	AGTGGTCAACCCAGGAGGTACTACGGGAGAGGCTCAACGCCATCATCGTCTTCGACG	538
QY	960	CTGCTTCTTCAGACAGCAGCTCCCGACTACCACTACATCATGAGGAACTCTTCT	1019
DB	539	CTGCTTCTTCAGACAGCAGCTCCCGACTACCACTACATCATGAGGAACTCTTCT	598
QY	1020	GTAAGTATCAGCAGCTTACGCTCTGTGGTGGAGTACATGATCGTGATGCTGAA	1079
DB	599	GTAAGTATCAGCAGCTTACGCTCTGTGGTGGAGTACATGATCGTGATGCTGAA	658
QY	1080	CGGTGCCACGCCCGGAGGATGCTGGATCGGCTGGAGAGATGCTACAGAT	1139
DB	659	CGGTGCCACGCCCGGAGGATGCTGGATCGGCTGGAGAGATGCTACAGAT	718
QY	1140	GATCGACCGGAGTTGGGAAAAAAGCTGAAGTCTTGTATCATGCTCCACCCCTGTGTT	1199
DB	719	GATCGGCGGAGTTGGGAAAAAAGCTGAAGTCTTGTATCATGCTCCACCCCTGTGTT	778
QY	1200	CATTGGGACCTGTGGGCACTCTTCGCCCTTTCATCAGCGTCAAGTTTCAACAAAGAT	1259
DB			
QY	779	CATTGGGACCTGTGGGCACTCTTCGCCCTTTCATCAGCGTCAAGTTTCAACAAAGAT	838
DB	1260	CCAGTACGTGCACAGCTTGGAGAGCTGGAGCAACTCATCCCTATGGACACAGTCCAGAT	1319
DB	839	CCAGTACGTGCACAGCTTGGAGAGCTGGAGCAACTCATCCCTATGGACACAGTCCAGAT	898
QY	1320	CCAGACTGCGTCTGCAATACGAAGAGGAAAGACTGAAGGCCAGAGGAGGAGCGCGAG	1379
DB	899	CCAGACTGCGTCTGCAATACGAAGAGGAAAGACTGAAGGCCAGAGGAGGAGCGCGAG	958
QY	1380	GGCCAGCGGAGTTGTGCTGCCAGGTCTGGAAGAGCCAGAGGTGGCAACAGTGA	1439
DB	959	GGCCAGCGGAGTTGTGCTGCCAGGTCTGGAAGAGCCAGAGGTGGCAACAGTGA	1018
QY	1440	AAACAGGTCTGCTGTGCTCTCAAGAGATCAGGAACAGCAATCTCTGAGGCGAGTGA	1499
DB	1019	AAACAGGTCTGCTGTGCTCTCAAGAGATCAGGAACAGCAATCTCTGAGGCGAGTGA	1078
QY	1500	CATAACAAAGGACATGGAAGAGATTCCAGATGCCAGAAAACTCTGTCAAGCGCCACT	1559
DB	1079	CATAACAAAGGACATGGAAGAGATTCCAGATGCCAGAAAACTCTGTCAAGCGCCACT	1138
QY	1560	GGCCCGAGATCTATCTGCTCTCATCTGAGTCCCAATCTTCAAGGGTGCAGCCCTTC	1619
DB	1139	GGCCCGAGATCTATCTGCTCTCATCTGAGTCCCAATCTTCAAGGGTGCAGCCCTTC	1198
QY	1620	CGTTCACTCTGGAACCCAGCATCTCTTTTCACTGCTTGAACATGTTATTTTTTTTT	1679
DB	1199	CGTTCACTCTGGAACCCAGCATCTCTTTTCACTGCTTGAACATGTTATTTTTTTTT	1258
QY	1680	TTAAGCATGAGTATTTGTGGTTCAGAAAAAGGCCAGCTCTGAGCCCTTCACCCCTTC	1739
DB	1259	TTAAGCATGAGTATTTGTGGTTCAGAAAAAGGCCAGCTCTGAGCCCTTCACCCCTTC	1318
QY	1740	CACACTCAAGAACTCTCAGCGAGGAGCAAGAACGCGGGGTGGCGGTGGCGT	1799
DB	1319	CACACTCAAGAACTCTCAGCGAGGAGCAAGAACGCGGGGTGGCGGTGGCGT	1378
QY	1800	CGGTGGCTCGCTCTGCTCGAGCCCTCTGTGTGAGAGGAGGAGGAGGAGGAGGAGG	1859
DB	1379	CGGTGGCTCGCTCTGCTCGAGCCCTCTGTGTGAGAGGAGGAGGAGGAGGAGGAGG	1438
QY	1860	CTTCTCTGTGCTGTGCTCAGCTCCAGGTCGAGGACACAGACCCAGCCGCCCGGCG	1919
DB	1439	CTTCTCTGTGCTGTGCTCAGCTCCAGGTCGAGGACACAGACCCAGCCGCCCGGCG	1498
QY	1920	TGGGTGCTGCTCTTCTGCTGCTTCTCCAGAGATGCGGCTCAGACCTTAGAGCTC	1979
DB	1499	TGGGTGCTGCTCTTCTGCTGCTTCTCCAGAGATGCGGCTCAGACCTTAGAGCTC	1558
QY	1980	AACCCCTATGAGGCGACCTCTGGGTAGCTCTGAGCTCCGACCTTAGTCCAAAT	2039
DB	1559	AACCCCTATGAGGCGACCTCTGGGTAGCTCTGAGCTCCGACCTTAGTCCAAAT	1618
QY	2040	TTCAACCCATGTTTTCATTTGACCGCCCTCTCTGCTCATATAAGACACCCAGCTC	2099
DB	1619	TTCAACCCATGTTTTCATTTGACCGCCCTCTCTGCTCATATAAGACACCCAGCTC	1678
QY	2100	CTTTGAGGATCAGAGCCCATTTGCAAGAGAGCGCTGCCAACCATCTGTCTCTCC	2159
DB	1679	CTTTGAGGATCAGAGCCCATTTGCAAGAGAGCGCTGCCAACCATCTGTCTCTCC	1738
QY	2160	GATTGCAAAATGACACCCAGTAACTTAGAAATTTCTCAAGCCCTTTAACTCAGATGC	2219
DB	1739	GATTGCAAAATGACACCCAGTAACTTAGAAATTTCTCAAGCCCTTTAACTCAGATGC	1798
QY	2220	AAGCCACCGGCAAAACCGTCAATCTCCCAAGAGGATGAGATGTGGACCTCAC	2279
DB	1799	AAGCCACCGGCAAAACCGTCAATCTCCCAAGAGGATGAGATGTGGACCTCAC	1858
QY	2280	TGCTCCCGCAACCCAGCGTGGGACACGCGCAACCGCTGTTCGGGGTGGAAACGCA	2339
DB	1859	TGCTCCCGCAACCCAGCGTGGGACACGCGCAACCGCTGTTCGGGGTGGAAACGCA	1918

599 GTACGTCATCAGCAGCTTACAGCTCTCTGGTGGCTGAGGACTACATGATCGTGACCTGAA 658
1080 CGGTGCCACGCCCGCGGAGGATGCTCGAATCGGCTGGCTGAAGAAGTGCTTACCAGAT 1139
659 CGGTGCCACGCCCGCGGAGGATGCTCGAATCGGCTGGCTGAAGAAGTGCTTACCAGAT 718
1140 GATCGACCGGAGGTGGGAAACCTGAAAGTCTCTTGATCATCGTCCACCCCTCGTGGT 1199
719 GATCGCCCGGAGGTGGGAAACCTGAAAGTCTCTTGATCATCGTCCACCCCTCGTGGT 778
1200 CATTCGGAATGCTGCGCCATCTCTCGCCCTTTCATCAGGCTCAAGTTCATCAAGAAT 1259
779 CATTCGGAATGCTGCGCCATCTCTCGCCCTTTCATCAGGCTCAAGTTCATCAAGAAT 838
1260 CCAGTACGTCACAGCTTGGAAACCTTGGACCACTCATCCCTATGGAACACGCTCCAGAT 1319
839 CCAGTACGTCACAGCTTGGAAACCTTGGACCACTCATCCCTATGGAACACGCTCCAGAT 898
1320 CCAGTACGTCCTGCAATACGAAGAGAAAGACTGAAGGCGAGGAGGAGCGCGAG 1379
899 CCAGTACGTCCTGCAATACGAAGAGAAAGACTGAAGGCGAGGAGGAGCGCGAG 958
1380 GCGCCAGCGGAGTTTGTGTCGCCAGGCTTGAAGAGAAAGCTGAGAGTGGCAGCAGTGGA 1439
959 GCGCCAGCGGAGTTTGTGTCGCCAGGCTTGAAGAGAAAGCTGAGAGTGGCAGCAGTGGA 1018
1440 AAACAGTCTGCTCTGCTCTCAGAGATCAGGAAACAGCATGCTCTGAGGCGAGCTGAG 1499
1019 AAACAGTCTGCTCTGCTCTCAGAGATCAGGAAACAGCATGCTCTGAGGCGAGCTGAG 1078
1500 CATACAAAGGACATGGAAGAGATTCAGATGCGCAAGAAACCTCTGTGAGAGCGCCACT 1559
1079 CATACAAAGGACATGGAAGAGATTCAGATGCGCAAGAAACCTCTGTGAGAGCGCCACT 1138
1560 GCGCCAGATCTCATCTGCTCTCATCTGAGTCCCAATCTTCAAGGCTGCGAGCGCCCTC 1619
1139 GCGCCAGATCTCATCTGCTCTCATCTGAGTCCCAATCTTCAAGGCTGCGAGCGCCCTC 1198
1620 CGTTCACTCTGAAACCCAGCATCTCTTTCAGTCTGCTGCAAAACATGTAATTTTTTTT 1679
1199 CGTTCACTCTGAAACCCAGCATCTCTTTCAGTCTGCTGCAAAACATGTAATTTTTTTT 1258
1680 TTAACGATCAGTATTTGTGCGTTCCAGAAAGGCGCCAGCTCTGAGCCCTCAGCCCTTC 1739
1259 TTAACGATCAGTATTTGTGCGTTCCAGAAAGGCGCCAGCTCTGAGCCCTCAGCCCTTC 1318
1740 CACACTCAGAACTCTCAGCGGAGGAGCAAGAGCGGAGGCGTGGCCCGGTGGCGT 1799
1319 CACACTCAGAACTCTCAGCGGAGGAGCAAGAGCGGAGGCGTGGCCCGGTGGCGT 1378
1800 CGGTGGCTCCGCTCTGCTCTGCGAGCCCTGCTGCTCAGAGCTGGATACAGATTCAGAC 1859
1379 CGGTGGCTCCGCTCTGCTCTGCGAGCCCTGCTGCTCAGAGCTGGATACAGATTCAGAC 1438
1860 CTTCTCTTGTGCTGTCACCCGCTTCCAGGTTGGAGCCACAGACACCCAGCCCGCGC 1919
1439 CTTCTCTTGTGTCACCCGCTTCCAGGTTGGAGCCACAGACACCCAGCCCGCGC 1498
1920 TGGGTCTGCTCTTCTGCTGCTTCCCTCAGAAAGCGGCTCAGACCTTAGAGCTC 1979
1499 TGGGTCTGCTCTTCTGCTGCTTCCCTCAGAAAGCGGCTCAGACCTTAGAGCTC 1558
1980 AACCCCTTATGAGGCGCAGCTCTGGGTAGCTCTGACCTCGACCTTATGTCAT 2039
1559 AACCCCTTATGAGGCGCAGCTCTGGGTAGCTCTGACCTCGACCTTATGTCAT 1618
2040 TTCACACCATGTTTTCATTTGACCCCGCCCTTCTCGCTCATATGACACCCAGCTC 2099
1619 TTCACACCATGTTTTCATTTGACCCCGCCCTTCTCGCTCATATGACACCCAGCTC 1678
2100 CTTTGAGGATCAGAGCCCATTTGACAAAGAGCGCTGCGCAACCATCTTGTCTCC 2159

1679 CTTTGAGGATCAGAGCCCATTTGCACAAAGAGCGCTGCGCAACCATCTTGTCTCTC 1738
2160 GATTGCAAAATCAGACCCCGAGTAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGC 2219
1739 GATTGCAAAATCAGACCCCGAGTAATCTAGAACATTTCTCAAGCCCTTTAACTCAGATGC 1798
2220 AAGCCACCGGGCAAAACCCCGTCAATCTCCCAACCAAGGAATGAGATATGTGGACCTCAC 2279
1799 AAGCCACCGGGCAAAACCCCGTCAATCTCCCAACCAAGGAATGAGATATGTGGACCTCAC 1858
2280 TGCTCTCCCAACCCCGAGCTCAGGCTGGGACACCGCAACGCTTCTCGGGTTGGAACAGCA 2339
1859 TGCTCTCCCAACCCCGAGCTCAGGCTGGGACATGCAACGCTTCTCGGGTTGGAACAGCA 1918
2340 GAGGCTCAGAACTCGCTCTGAAATAGGACAGCTTAGCAAGAGGAAGATACAGGGTATCG 2399
1919 GAGGCTCAGAACTCGCTCTGAAATAGGACAGCTTAGCAAGAGGAAGATACAGGGTATCG 1978
2400 GCGGTTGAGTGTTCAGAGTCAATTCGGGAAGATAAATCCAGTGGCTGGCGCGAGCCA 2459
1979 GCGGTTGAGTGTTCAGAGTCAATTCGGGAAGATAAATCCAGTGGCTGGCGCGAGCCA 2038
2460 CTTGATTCAGAACTTGGACACGCGGTTCTTGTTCGGGAGCAAAATTTCCCTAGAAAA 2519
2039 CTTGATTCAGAACTTGGACACGCGGTTCTTGTTCGGGAGCAAAATTTCCCTAGAAAA 2098
2520 AGAAGACAGACTTTTCTAATGTGTCCAATGCGGATCACTGCTCAGATGGACTTAGAA 2579
2099 AGAAGACAGACTTTTCTAATGTGTCCAATGCGGATCACTGCTCAGATGGACTTAGAA 2158
2580 GCACGTGAGCTCCCTGCTCTGGAAGTATTTAAGAAAAAGCTGGGCGAGCAGATGGCTC 2639
2159 GCACGTGAGCTCCCTGCTCTGGAAGTATTTAAGAAAAAGCTGGGCGAGCAGATGGCTC 2218
2640 AGCCTGTAACTCCAGACTTTGGGAGGCGGAGGAGGAGGATCACTGAGGTGAGGAGTT 2699
2219 AGCCTGTAACTCCAGACTTTGGGAGGCGGAGGAGGAGGATCACTGAGGTGAGGAGTT 2278
2700 TGAGAACAGCTGGCCCAACATGTTGAAACCTCTCTACTTAAATAACAAAAATTAGCC 2759
2279 TGAGAACAGCTGGCCCAACATGTTGAAACCTCTCTACTTAAATAACAAAAATTAGCC 2338
2760 AGCGTGTGGCAGGCTGCTGTAATCCAGCTACTTTGGAGGCTCAGGCGATGAGATCAC 2819
2339 AGCGTGTGGCAGGCTGCTGTAATCCAGCTACTTTGGAGGCTCAGGCGATGAGATCAC 2398
2820 TTAACCTGAGAGGAGGTTACAGTACGAGCAAGATCGTGCACCTGCAATTCAGCCTGG 2879
2399 TTAACCTGAGAGGAGGTTACAGTACGAGCAAGATCGTGCACCTGCAATTCAGCCTGG 2458
2880 GCGACAGAGCAAGCTCTCTCAAAAAA 2917
2459 GCGACAGAGCAAGCTCTCTCAAAAAA 2496

RESULT 14
BC008736
LOCUSDEFINITION
Homo sapiens ataxia, cerebellar, Cayman type (caytaxin), mRNA (CDNA
clone IMAGE:3162511), partial cds.ACCESSION
BC008736VERSION
BC008736.2 GI:38014014KEYWORDS
SOURCEORGANISM
Homo sapiens (human)REFERENCE
AUTHORS

BC008736 2276 bp mRNA linear PRI 12-NOV-2003
Homo sapiens ataxia, cerebellar, Cayman type (caytaxin), mRNA (CDNA
clone IMAGE:3162511), partial cds.
BC008736
BC008736.2 GI:38014014
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo
1 (bases 1 to 2276)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Young, A.C., Shevchenko, Y., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Kravinsky, M.I., Skalska, U., Smal, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2276)

Strausberg, R.

Direct Submission

Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Oct 28, 2003 this sequence version replaced gi:14250564.

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mascian, S.D., McIlone, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>

Series: IRAL Plate: 5 Row: 0 Column: 7

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis.

Location/Qualifiers

1..2276

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3162511"

/issue_type="Brain, neuroblastoma"

/clone_lib="NIH_MGC_19"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

ORIGIN

Query Match 62.2%; Score 1815.8; DB 8; Length 2276;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1902; Conservative 0; Mismatches 2; Indels 65; Gaps 2;

QY 1013 TCTTCTGTACGTCATCAGCAGCTTAGAGCTCTGGTGGCTGAGGACTACATGATCGTGT 1072
|||||
DB 1 TCTTCTGTACGTCATCAGCAGCTTAGAGCTCTGGTGGCTGAGGACTACATGATCGTGT 60
|||||

QY 1073 ACCTGAACGGTGCACGCCCGCGGAGGATCGCTGGAATCGGCTGGTGAAGAGATGCT 1132
|||||

Db 61 ACCTGAACGGTGCACGCCCGCGGAGGATCGCTGGAATCGGCTGGTGAAGAGATGCT 120
|||||
QY 1133 ACCAGATCATCGACCGAGGTTGGGGAACCTGAGTCTCTTGAATCATCGTCCACCCCT 1192
|||||
Db 121 ACCAGATCATCGACCGAGGTTGGGGAACCTGAGTCTCTTGAATCATCGTCCACCCCT 180
|||||
QY 1193 COTGGTTCATTCGGGACTGTGTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCATCA 1252
|||||
Db 181 COTGGTTCATTCGGGACTGTGTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCATCA 240
|||||
QY 1253 ACAAGATCCAGTACGTGACAGCTTGGAGACCTTGGAGCAACTATCCCTATGGAACAG 1312
|||||
Db 241 ACAAGATCCAGTACGTGACAGCTTGGAGACCTTGGAGCAACTATCCCTATGGAACAG 300
|||||
QY 1313 TCCAGATCCAGACTGCGTCTGCAATACGAGGAGGAGAGCTGAGGCCAGGAGGAGAGA 1372
|||||
Db 301 TCCAGATCCAGACTGCGTCTGCAATACGAGGAGGAGAGCTGAGGCCAGGAGGAGAGA 360
|||||
QY 1373 GCGGAGGCCCGCAGCGGAGGTTTGTGTGCGCCAGGTCTGGAAGAGAGCCAGAGGTGGCAC 1432
|||||
Db 361 GCGGAGGCCCGCAGCGGAGGTTTGTGTGCGCCAGGTCTGGAAGAGAGCCAGAGGTGGCAC 420
|||||
QY 1433 CAGTGGAAAAAGAGTCTGTCTGTCTCAGAGATCAGGAAACAAG----- 1478
|||||
Db 421 CAGTGGAAAAAGAGTCTGTCTGTCTCAGAGATCAGGAAACAAGCCCTCCAGCCCTGC 480
|||||
QY 1479 -----CATGTCCTGA 1488
|||||
Db 481 ACAGATCAAGTGTATGCCACGTGGCCCAAGTTTCAACCTAAGTCACTGTCCTGA 540
|||||
QY 1489 GCGGAGGTGAGCAATAACAAAGGACATGGAAGAGATTCAGATGCCAGAAACCTCTGTC 1548
|||||
Db 541 GCGGAGGTGAGCAATAACAAAGGACATGGAAGAGATTCAGATGCCAGAAACCTCTGTC 600
|||||
QY 1549 AGAGCCCACTGGGCCCGCAGATCTCATCTGCTCTCATCTGAGTCCCAATTTTCAAGGT 1608
|||||
Db 601 AGAGCCCACTGGGCCCGCAGATCTCATCTGCTCTCATCTGAGTCCCAATTTTCAAGGT 660
|||||
QY 1609 GCCAGCCCTCCGTTTCATCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAACAATGT 1668
|||||
Db 661 GCCAGCCCTCCGTTTCATCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAACAATGT 720
|||||
QY 1669 ATTTTTCATTCAGTATGAGTATTTGTGCGTTTTCAGAAAGGCGCCAGCTCTGAGCC 1728
|||||
Db 721 A-TTTTTCATTCAGTATGAGTATTTGTGCGTTTTCAGAAAGGCGCCAGCTCTGAGCC 779
|||||
QY 1729 CCTCACCCTTCCACACTCAGAACTCTCAGCCGAGGAGGCAAGAGCGAGGGGTGGC 1788
|||||
Db 780 CCTCACCCTTCCACACTCAGAACTCTCAGCCGAGGAGGCAAGAGCGAGGGGTGGC 839
|||||
QY 1789 CCGGTGGCGTGGTGGCTCGCTCTGCTGCGAGCCCTGTGTGCTGAGCTGATACA 1848
|||||
Db 840 CCGGTGGCGTGGTGGCTCGCTCTGCTGCGAGCCCTGTGTGCTGAGCTGATACA 899
|||||
QY 1849 AGATTCAAGACCTTCTCTGCTGTGTACCCGCTCCAGGTGGAGCCACAGACCCACC 1908
|||||
Db 900 AGATTCAAGACCTTCTCTGCTGTGTACCCGCTCCAGGTGGAGCCACAGACCCACC 959
|||||
QY 1909 GCCACCCCGGTGGTGGTGGTCTCTTCTGCTGCTTTTCCCTTCCAGAAATGGGGCTCAGA 1968
|||||
Db 960 GCCACCCCGGTGGTGGTGGTCTCTTCTGCTGCTTTTCCCTTCCAGAAATGGGGCTCAGA 1019
|||||
QY 1969 CCTAGAGCTCAACCCCTTATGAGGCGCAGTCTCTGGGTAGTCTGAGCTCCGAGCT 2028
|||||
Db 1020 CCTAGAGCTCAACCCCTTATGAGGCGCAGTCTCTGGGTAGTCTTCTGAGCTCCGAGCT 1079
|||||
QY 2029 TATGTCCAAATTTTCAACCCCTTATGAGGCGCAGTCTCTGGGTAGTCTTCTGCTCATATG 2088
|||||
Db 1080 TATGTCCAAATTTTCAACCCCTTATGAGGCGCAGTCTCTGGGTAGTCTTCTGCTCATATG 1139
|||||
QY 2089 ACACCCAGCTCTTTGAGAGGATCAGAGCCCATTTGCAAGAGAGCGCTGCCAACAT 2148
|||||

FEATURES

source

Db 1140 ACACCCAGCTCCTTTGAGGATCAGAGCCATTGTCACAGAGAGCGCGTGCACACCAT 1199
QY 2149 CTTGTGCTCCGATGCAAAATGACACCCAGTAATCTAGAAATCTCAAGCCCTTTA 2208
Db 1200 CTTGTGCTCCGATGCAAAATGACACCCAGTAATCTAGAAATCTCAAGCCCTTTA 1259
QY 2209 ACTCAGATGTCAGACCCAGCGGCAAAACCCGCTCAATCTCCACCAAGATGAGATAT 2268
Db 1260 ACTCAGATGTCAGACCCAGCGGCAAAACCCGCTCAATCTCCACCAAGATGAGATAT 1319
QY 2269 GTGAGCTCACTGCTCCCCCAACCCAGCGTTCAGCTGGGACCGCAACCGCTGTCGGG 2328
Db 1320 GTGAGCTCACTGCTCCCCCAACCCAGCGTTCAGCTGGGACCGCAACCGCTGTCGGG 1379
QY 2329 TTGGAACAGCAGAGGCTCAGAACTGGCTCTGAAATAGGACAGCTAGCAAGAGGAAGAT 2388
Db 1380 TTGGAACAGCAGAGGCTCAGAACTGGCTCTGAAATAGGACAGCTAGCAAGAGGAAGAT 1439
QY 2389 ACAGGATATCGGCGTTTGAAGTCTTTCAGAGTCAATTCGGGAAGATAAATCCAGTGCCT 2448
Db 1440 ACAGGATATCGGCGTTTGAAGTCTTTCAGAGTCAATTCGGGAAGATAAATCCAGTGCCT 1499
QY 2449 GCGCGCAGCACCCTGCAATCAAGCTTGGACAGCGGTTCTGTTGGGAGGCAAAATTT 2508
Db 1500 GCGCGCAGCACCCTGCAATCAAGCTTGGACAGCGGTTCTGTTGGGAGGCAAAATTT 1559
QY 2509 CCTAGGAAAAAGAACAGACAGCTTTTCTAATGTGTCCAAATCGCGATCACTGGTCAGAT 2568
Db 1560 CCTAGGAAAAAGAACAGACAGCTTTTCTAATGTGTCCAAATCGCGATCACTGGTCAGAT 1619
QY 2569 GGAATCTAGAGCAGTGAAGTCTCTCTCTGGAAGTATTAAAGAAAGGCTGGGCGAG 2628
Db 1620 GGAATCTAGAGCAGTGAAGTCTCTCTCTGGAAGTATTAAAGAAAGGCTGGGCGAG 1679
QY 2629 CAGATGGCTCAGCGCTGTAATCCAGACTTTGGAGGCGGAGGAGCGGATCACTGA 2688
Db 1680 CAGATGGCTCAGCGCTGTAATCCAGACTTTGGAGGCGGAGGAGCGGATCACTGA 1739
QY 2689 GGTGAGGAGTTTGAAGAACAGCTGGCCAAATGTTGAAACCTCATCTCTACTAAATAATAC 2748
Db 1740 GGTGAGGAGTTTGAAGAACAGCTGGCCAAATGTTGAAACCTCATCTCTACTAAATAATAC 1799
QY 2749 AAAAAATTAGCAGCGTGTGGCAGTGCTGTAAATCCAGTACTTTGGAGGCTGAGGC 2808
Db 1800 AAAAAATTAGCAGCGTGTGGCAGTGCTGTAAATCCAGTACTTTGGAGGCTGAGGC 1859
QY 2809 ATGAGATCACTTAACTGAGGAGGAGGTTACAGTGGCCAGATGTCGCCACTGA 2868
Db 1860 ATGAGATCACTTAACTGAGGAGGAGGTTACAGTGGCCAGATGTCGCCACTGA 1919
QY 2869 TTCCAGCTGGGCGCAGAGCAAGACTCTGTCTCAAAAAAATAAAAAA 2917
Db 1920 TTCCAGCTGGGCGCAGAGCAAGACTCTGTCTCAAAAAAATAAAAAA 1968

RESULT 15
AC011488
LOCUS AC011488 134308 bp DNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens chromosome 19 clone CTB-171N13, complete sequence.
ACCESSION AC011488
VERSION AC011488.7 GI:22038500
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 134308)
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 134308)
AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 134308)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL Direct Submission
TITLE Submitted (01-AUG-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Aug 1, 2002 this sequence version replaced gi:9256301.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
FEATURES
source
1..134308
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-171N13"
ORIGIN
Query Match 49.3%; Score 1438.2; DB 8; Length 134308;
Best Local Similarity 99.8%; Pred. No. 2.8e-272;
Matches 1440; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1475 CAAGATGTCCTGAGGCGACGTGAGCATAAACAGGACATGGAAGAATCCAGATGCC 1534
Db 74150 CTAGCATGTCCTGAGGCGACGTGAGCATAAACAGGACATGGAAGAATCCAGATGCC 74209
QY 1535 AGAAACCTCTGTCAGAGCCCACTGGCCAGATCTCATCTGCTCATCTGAGTCCC 1594
Db 74210 AGAAACCTCTGTCAGAGCCCACTGGCCAGATCTCATCTGCTCATCTGAGTCCC 74269
QY 1595 AATCTTCAAGGGTCCAGCCCTCCGTTTCATCTCTGAAACCCAGCATCTTTTCAGCTG 1654
Db 74270 AATCTTCAAGGGTCCAGCCCTCCGTTTCATCTCTGAAACCCAGCATCTTTTCAGCTG 74329
QY 1655 CTTGAAAAATGTAATTTTTTTTTTAAACGATGAGTATTGTGCGTTCCAGAAAAAGG 1714
Db 74330 CTTGAAAAATGTAATTTTTTTTTTAAACGATGAGTATTGTGCGTTCCAGAAAAAGG 74389
QY 1715 CCCAGCTCTGAGCCCTCACCCCTTCCACACTCAGAACTCTCAGCCGAGGAAGCAAGAA 1774
Db 74390 CCCAGCTCTGAGCCCTCACCCCTTCCACACTCAGAACTCTCAGCCGAGGAAGCAAGAA 74449
QY 1775 GCGCAGGGGGTGGCCGGTGGCGTGGCGTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 1834
Db 74450 GCGCAGGGGGTGGCCGGTGGCGTGGCGTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 74509
QY 1835 CAGAGCTGGATACAGATTAAGAGCCCTTCTCTTGTGTGTCAACCCGCTCCAGGTTGGAGC 1894
Db 74510 CAGAGCTGGATACAGATTAAGAGCCCTTCTCTTGTGTGTCAACCCGCTCCAGGTTGGAGC 74569
QY 1895 CACAGACACCAACCGCCACCCCGGTGGGTCTGCTCTCTCTGCTGCTCTCTCTCTCTCTC 1954
Db 74570 CACAGACACCAACCGCCACCCCGGTGGGTCTGCTCTCTCTGCTGCTCTCTCTCTCTCTC 74629
QY 1955 AATGGGCTCAGACCTAGAGCTCAACCCCGCTTATGAGGCGCACGCTCTGGGGTAGCTC 2014
Db 74630 AATGGGCTCAGACCTAGAGCTCAACCCCGCTTATGAGGCGCACGCTCTGGGGTAGCTC 74689
QY 2015 CTGACCTCCGACCTTATGTCCAAATTTTCAACCCATGTTTTTCAATTTGACCCGCCCT 2074
Db 74690 CTGACCTCCGACCTTATGTCCAAATTTTCAACCCATGTTTTTCAATTTGACCCGCCCT 74749
QY 2075 TCTGCTCATATGACACCCAGCTCTCTTTGAGAGATCAGAGCCCATGTCACAGAGAG 2134
Db 74750 TCTGCTCATATGACACCCAGCTCTCTTTGAGAGATCAGAGCCCATGTCACAGAGAG 74809
QY 2135 CCGCTGCCAACCATCTTGTCTCCGATTGCAAAATGACACCCCGCATCTAGAACATT 2194

```
||||| 74810 CCCTGCCAACCATCTTGTCTCCGATTGCAAAATGACACCCAGTAATCTAGAACATT 74869
QY 2195 CTCAGCCCCCTTTAACTCAGATGTCAGCCACCGGCAACCCCGTCAATACCTCCACC 2254
Db 74870 CTCAGCCCCCTTTAACTCAGATGTCAGCCACCGGCAACCCCGTCAATACCTCCACC 74929
QY 2255 AAGGAATGAGATATGTGGACCTCACTGCTCCCCCAACCCAGCGTCAGGCTGGGACACGCC 2314
Db 74930 AAGGAATGAGATATGTGGACCTCACTGCTCCCCCAACCCAGCGTCAGGCTGGGACACGCC 74989
QY 2315 AAGCTGTTCCGGGTTGGAAACAGCAGAGGCTCAGAACTGGCTCTGAAATAGGACACCT 2374
Db 74990 AAGCTGTTCCGGGTTGGAAACAGCAGAGGCTCAGAACTGGCTCTGAAATAGGACACCT 75049
QY 2375 AGCAAGAGGAAGATACAGGGTATCGGGCGTTTGAGTGTTCAGAACTCATTCGGGAAGAT 2434
Db 75050 AGCAAGAGGAAGATACAGGGTATCGGGCGTTTGAGTGTTCAGAACTCATTCGGGAAGAT 75109
QY 2435 AAATCCAGTGCCTGGCCGAGCACCTTGCATTCAAAGCTTGACCCAGCGGTTCTTGT 2494
Db 75110 AAATCCAGTGCCTGGCCGAGCACCTTGCATTCAAAGCTTGACCCAGCGGTTCTTGT 75169
QY 2495 CGGAGGCAAAATTCCTTAGGAAAGAAAGACAGACTTTTCTAATGTGTCCAAATGCGG 2554
Db 75170 CGGAGGCAAAATTCCTTAGGAAAGAAAGACAGACTTTTCTAATGTGTCCAAATGCGG 75229
QY 2555 ATCACTGGTCAGATGGACTCTAGAAGCACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAA 2614
Db 75230 ATCACTGGTCAGATGGACTCTAGAAGCACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAA 75289
QY 2615 AAGGCTGGGCCAGGCACGATGGCTCACGGCTGTAAATCCAGACTTTGGGAGGCCGAGGCA 2674
Db 75290 AAGGCTGGGCCAGGCACGATGGCTCACGGCTGTAAATCCAGACTTTGGGAGGCCGAGGCA 75349
QY 2675 GCGGATCACTGAGTGAGGAGTTTGAGAAACAGCCTGGCCAAACATGGTGAAACCTCATC 2734
Db 75350 GCGGATCACTGAGTGAGGAGTTTGAGAAACAGCCTGGCCAAACATGGTGAAACCTCATC 75409
QY 2735 TCTACTAAAAATACAAAAATTAGCCAGGCGTGGTGGCAGGTGCTGTATCCCGCTACT 2794
Db 75410 TCTACTAAAAATACAAAAATTAGCCAGGCGTGGTGGCAGGTGCTGTATCCCGCTACT 75469
QY 2795 TGGAGGCTGAGGCATGAGAAATCACTTAAACCTGAGAGGCAGAGGTTACAGTGAGCCAAAG 2854
Db 75470 TGGAGGCTGAGGCATGAGAAATCACTTAAACCTGAGAGGCAGAGGTTACAGTGAGCCAAAG 75529
QY 2855 ATCGTCCCACTGCATTCAGCTTGGGCGACAGAGCAAGACTCTGTCTCAAAAAAATAA 2914
Db 75530 ATCGTCCCACTGCATTCAGCTTGGGCGACAGAGCAAGACTCTGTCTCAAAAAAATAA 75589
QY 2915 AAA 2917
Db 75590 AAA 75592
```

Search completed: March 5, 2006, 07:27:41
Job time : 14453 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 02:57:19 ; Search time 1671 Seconds
(without alignments)

11638.284 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gccgagcctctgcagccct.....tctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002s:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003s:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004s:*
- 13: Geneseq2004bs:*
- 14: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2918	100.0	2918	13	ADQ16276 Human Cay
2	2916.4	99.9	2918	13	ADQ16281 Human Cay
3	2913.8	99.9	4978	12	ADQ85643 Human tum
4	2913.8	99.9	4978	13	ADQ40413 Human tum
5	2896.6	99.3	4976	6	ADQ83973 Human gen
6	2887.6	99.0	3065	3	ADQ00128 Human NIP
7	2870.2	98.4	2894	4	AAH16751 Human cDN
8	2598.2	80.0	2759	10	ADQ62843 Human cDN
9	2357.6	80.8	2786	12	ADQ63969 Novel hum
10	1438.2	49.3	47219	13	ADQ16283 Human Cay
11	1438.2	49.3	59884	13	ADQ16284 Human Cay
12	1325.8	45.4	2625	10	ADQ62832 Human cDN
13	854.6	29.3	2166	13	ADQ16274 Mouse jit
14	745.2	25.5	790	4	AAH08588 Human cDN
15	675	23.1	703	5	AAH80323 DNA encod
16	579	19.8	579	12	ADQ92257 Human aut
17	470	16.1	560	4	AAH13509 Human cDN
18	420.4	14.4	485	9	AAH15328 Human adu
19	331.6	11.4	384	2	AAQ59527 Human bra

20	330.4	11.3	5656	12	ADQ21252 Human sof
21	327.6	11.2	5668	13	ADQ82885 Human KIA
22	327.6	11.2	11690	9	ACC45236 Human BMC
23	318.2	10.9	5654	6	ABL62307 Colon ade
24	318.2	10.9	5654	6	ABL68862 Kidney ca
25	318.2	10.9	5654	6	ABL68862 Kidney ca
26	315.8	10.8	2347	4	AAI98076 Human neu
27	302.2	10.4	4235	3	AAAD00129 Human NTP
28	302.2	10.4	4378	5	ABV27987 Human pro
29	302.2	10.4	4378	5	ABV22147 Human pro
30	259.2	8.9	273	5	ABAI3564 Human ner
31	245	8.4	2417	13	ACN40083 Tumour-as
32	243.6	8.3	660	5	ADL39125 Human ova
33	243.4	8.3	2544	4	AAH14059 Human cDN
34	242.2	8.3	139308	8	ADB12769 Human PRK
35	241.8	8.3	96593	11	ADL27146 Human gen
36	241.8	8.3	96595	9	ADA03068 Human PPP
37	241.8	8.3	96595	10	ADB72806 Human PPP
38	241.8	8.3	96596	9	ADA66352 Human PPP
39	241.6	8.3	110000	10	ADE1169 2 Continuation (3 of
40	240.8	8.3	138627	12	ADQ97183 Human can
41	240	8.2	100608	13	ABD32769 Human can
42	239.6	8.2	20001	13	ADT77129 Type II d
43	239.6	8.2	32249	4	AAI05336 Human rep
44	239.6	8.2	32249	4	ABL98205 Human tes
45	239.2	8.2	8939	5	ABAI8915 Human ner

ALIGNMENTS

RESULT 1

ADQ16276

ID ADR16276 standard; cDNA; 2918 BP.

XX ADR16276;

XX ADR16276;

XX 21-OCT-2004 (first entry)

XX Human Cayman ataxia cDNA.

XX Human; jittery; ss; gene; Cayman ataxia; ATCAY; KIAA1872;

KW Chromosome 19p13.3; ataxia; myoclonus; dystonia; epileps; myastagnus;

KW SNP; single nucleotide polymorphism.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

XX CDS

XX 373..1488

XX /tag= a

XX /product= "Cayman ataxia protein"

XX replace(1275,G)

XX /tag= b

XX /standard_name= "Single nucleotide polymorphism"

XX US2004146900-A1.

XX 29-JUL-2004.

XX 03-NOV-2003; 2003US-00699941.

XX 01-NOV-2002; 2002US-0422971P.

XX 08-NOV-2002; 2002US-0424973P.

XX (UNMI) UNIV MICHIGAN.

XX Burmeister M;

XX WPI; 2004-552665/53.

XX P-PSDB; ADR16277.

XX Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in

XX subject, by detecting presence or absence of variant Cayman ataxia

XX PT

PT polypeptide or nucleic acid in biological sample.

XX Claim 4; SEQ ID NO 3; 98pp; English.

XX The invention relates to detecting variant Cayman ataxia polypeptide or
CC nucleic acid sequence (ATCAY) in a subject, involving providing a
CC biological sample from a subject, where the biological sample comprises a
CC Cayman ataxia polypeptide or nucleic acid, and detecting the presence or
CC absence of a variant Cayman ataxia polypeptide or nucleic acid in the
CC biological sample. Also included is a kit comprising a reagent for
CC detecting the presence or absence of a variant Cayman ataxia nucleic acid
CC or polypeptide in a biological sample. The human ATCAY gene (previously
CC isolated as cDNA KIA1872, located on chromosome 19p13.3) was identified
CC and mapped using the fact that it is the homologue of the mouse Jittery
CC gene, located on mouse chromosome 10. The method is useful for detecting
CC variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.
CC The presence of the variant Cayman ataxia polypeptide or nucleic acid is
CC indicative of a disorder chosen from ataxia, myoclonus, dystonia,
CC epilepsy, and nystagmus in the subject. The biological sample is chosen
CC from blood sample, a tissue sample, urine sample, saliva sample, and an
CC amniotic fluid sample. The subject is chosen from embryo, foetus, newborn
CC animal, young animal, and an adult animal. The animal is a human. The
CC human is an adult female of child-bearing age. The present sequence is
CC the cDNA encoding the Human ATCAY protein, the allele encoding the
CC normal, non-disease causing protein.

XX SQ Sequence 2918 BP; 699 A; 868 C; 797 G; 554 T; 0 U; 0 Other;

Query Match 100.0%; Score 2918; DB 13; Length 2918;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGAGCCTCTGCGAGCCTCTGAGCTGGGAAAGAGAGCTACCTCGGAGGACGGCGCGCA 60
DB 1 GCGGAGCCTCTGCGAGCCTCTGAGCTGGGAAAGAGAGCTACCTCGGAGGACGGCGCGCA 60
QY 61 GCGGCGCGCGATGAGAGGGGGCGAGCGGAGCCCGCGCTGGGGAGCCCAAC 120
DB 61 GCGGCGCGCGATGAGAGGGGGCGAGCGGAGCCCGCGCTGGGGAGCCCAAC 120
QY 121 CTGTGACCCACCCCTGACAAAGAGCTGGGGGGCTGGCCAGTCCCTTGG 180
DB 121 CTGTGACCCACCCCTGACAAAGAGCTGGGGGGCTGGCCAGTCCCTTGG 180
QY 181 TGACCTTCTCGGATGAGAGTCCGCTCGGAGCATCTCTTCTCTAGGCTTGA 240
DB 181 TGACCTTCTCGGATGAGAGTCCGCTCGGAGCATCTCTTCTCTAGGCTTGA 240
QY 241 GCGCGGGAGCGTGAGCGATGCGGCTGCGAGTCCCGGGGAGGGCTGTTTGGCA 300
DB 241 GCGCGGGAGCGTGAGCGATGCGGCTGCGAGTCCCGGGGAGGGCTGTTTGGCA 300
QY 301 GTAAGGAGGAGGCTGCTCAGCTGAGGGGGTCCCTGCTTCAAGCCAGTGCCTC 360
DB 301 GTAAGGAGGAGGCTGCTCAGCTGAGGGGGTCCCTGCTTCAAGCCAGTGCCTC 360
QY 361 TTCCAGCTCCCATGGGACCCAGGAGCAGCTCCGATGGAACCGTGAGGAG 420
DB 361 TTCCAGCTCCCATGGGACCCAGGAGCAGCTCCGATGGAACCGTGAGGAG 420
QY 421 GAGGAATGGCAGGACGAAAGATTTTCCAGGCCATCTCCAGAGAGAGCGGGTGGAACTG 480
DB 421 GAGGAATGGCAGGACGAAAGATTTTCCAGGCCATCTCCAGAGAGAGCGGGTGGAACTG 480
QY 481 CTGGCAGCCCGGTGGAAGACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCG 540
DB 481 CTGGCAGCCCGGTGGAAGACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCG 540
QY 541 CATCGTAAGAGGAGAGCTGGTGGCCCGCAGAGATCAACATTTCTCGGATCAGAGTGAG 600
DB 541 CATCGTAAGAGGAGAGCTGGTGGCCCGCAGAGATCAACATTTCTCGGATCAGAGTGAG 600
QY 601 GGGTCCCTGCTGCGATGACTTCTTTGGATACCCCTGATGACCTGGATATTAAACGTGGAT 660

DB 601 GGGTCCCTGCTGCTGCGATGACTTCTTTGGATACCCCTGATGACCTGGATATTAAACGTGGAT 660
QY 661 GACATCGAGAGCCCGGATGAGACCGACTCGCTGGAGTTCTTGGGGAATGGCAACCACTG 720
DB 661 GACATCGAGAGCCCGGATGAGACCGACTCGCTGGAGTTCTTGGGGAATGGCAACCACTG 720
QY 721 GAGTGGGAGAGCAGACACCCCGTGGCCACCGCAAGAAACATGCCCGGGGACAGCGCGAT 780
DB 721 GAGTGGGAGAGCAGACACCCCGTGGCCACCGCAAGAAACATGCCCGGGGACAGCGCGAT 780
QY 781 CTATTTGGGAGCGGACGACGCGAGGAGCGCGCGCCCAACCGGCGCTGTGGCGGACA 840
DB 781 CTATTTGGGAGCGGACGACGCGAGGAGCGCGCGCCCAACCGGCGCTGTGGCGGACA 840
QY 841 GTGATCATCGGGGAGCAAGAGCGATATAGACTTGCACATGATCCGGCCTTACATGAAA 900
DB 841 GTGATCATCGGGGAGCAAGAGCGATATAGACTTGCACATGATCCGGCCTTACATGAAA 900
QY 901 GTGGTCACTCCACGAGGAGGTACTTACGGCGAAGGCGCTCAACGGCATCATCGTCTTCAGCC 960
DB 901 GTGGTCACTCCACGAGGAGGTACTTACGGCGAAGGCGCTCAACGGCATCATCGTCTTCAGCC 960
QY 961 TGCTTCTTCAGAGCAGCAGCTCCCGACTACCACTACATCATGAGAACTCTTCTCTG 1020
DB 961 TGCTTCTTCAGAGCAGCAGCTCCCGACTACCACTACATCATGAGAACTCTTCTCTG 1020
QY 1021 TACGTCATCAGCGCTTACAGCTCTGCTGGCTGAGGACTACATGATCGTGTACCTGAAAC 1080
DB 1021 TACGTCATCAGCAGCTTACAGCTCTGCTGGCTGAGGACTACATGATCGTGTACCTGAAAC 1080
QY 1081 GGTGCCACCGCCCGGAGGAGTCTCGAATCGGCTGGCTGAAAGTGTCTACCAAGATG 1140
DB 1081 GGTGCCACCGCCCGGAGGAGTCTCGAATCGGCTGGCTGAAAGTGTCTACCAAGATG 1140
QY 1141 ATCGACCGGAGGTTGCGGAAAAAAGCTGAACTCTTGTATCATCGTCCACCTCTGCTGTT 1200
DB 1141 ATCGACCGGAGGTTGCGGAAAAAAGCTGAACTCTTGTATCATCGTCCACCTCTGCTGTT 1200
QY 1201 ATTGGAGCTGTGCTGGCCATCTCTGCGCTTTCATCAGGGTCAAGTTTATCAACAGATC 1260
DB 1201 ATTGGAGCTGTGCTGGCCATCTCTGCGCTTTCATCAGGGTCAAGTTTATCAACAGATC 1260
QY 1261 CAGTACGTGACAGCTTGGAGAGACCTGGAGCAACTCATCTTATGGAACAGTCCAGATC 1320
DB 1261 CAGTACGTGACAGCTTGGAGAGACCTGGAGCAACTCATCTTATGGAACAGTCCAGATC 1320
QY 1321 CCAGACTGCTGCTGCAATACGAGAGGAAAGACTGAAAGGCGGAGGAGGAGCGGAGG 1380
DB 1321 CCAGACTGCTGCTGCAATACGAGAGGAAAGACTGAAAGGCGGAGGAGGAGCGGAGG 1380
QY 1381 CCCAGCGGAGTTGTGCTGCTGCCAGGTTCTGAAAGAGAGCCAGAGTGGCACTGAGAA 1440
DB 1381 CCCAGCGGAGTTGTGCTGCTGCCAGGTTCTGAAAGAGAGCCAGAGTGGCACTGAGAA 1440
QY 1441 AACAGGCTGTCTGCTGCTCAGAGAGTCAAGAGAAACAGCATGTCTCAGGCGGAGTGGAGC 1500
DB 1441 AACAGGCTGTCTGCTGCTCAGAGAGTCAAGAGAAACAGCATGTCTCAGGCGGAGTGGAGC 1500
QY 1501 ATAAAGAGGACATGGAAGAGATTTCCAGATGCGCAAGAAACCTCTGTGAGAGCGCCACTG 1560
DB 1501 ATAAAGAGGACATGGAAGAGATTTCCAGATGCGCAAGAAACCTCTGTGAGAGCGCCACTG 1560
QY 1561 GCCCAGAGTCTATCTGCTGCTCATCTCAGTCCCAATCTTCCAGGGTGGCAGCGCCCTCC 1620
DB 1561 GCCCAGAGTCTATCTGCTGCTCATCTCAGTCCCAATCTTCCAGGGTGGCAGCGCCCTCC 1620
QY 1621 GTTCACTCTGAAACCCAGCATCTCTTTTTCAGTGTCTGAAACAACTGTTATTTTTTTTTT 1680
DB 1621 GTTCACTCTGAAACCCAGCATCTCTTTTTCAGTGTCTGAAACAACTGTTATTTTTTTTTT 1680
QY 1681 TAAAGATGAGATTTTGTGGGTTCCAGAAAGGGGCCAGTCTGAGCGCCCTCACCCCTTCC 1740

```
Db 1681 TAACGATCAGTATTTGTGGTTCAGAAAGGGCCAGCTCTGAGCCCTCCACCCCTTCC 1740
QY 1741 ACACTCAGAACTCTCAGCCGAGGAGCAAGAGCGAGGGGGTGGCCCGGCTGCGCTC 1800
Db 1741 ACACTCAGAACTCTCAGCCGAGGAGCAAGAGCGAGGGGGTGGCCCGGCTGCGCTC 1800
QY 1801 GGTGGCTCGCTCTGCTGCGAGCCCTGTGTGTGAGCTGGAATACAGATTCAGAGCC 1860
Db 1801 GGTGGCTCGCTCTGCTGCGAGCCCTGTGTGTGAGCTGGAATACAGATTCAGAGCC 1860
QY 1861 CTTCTCTCTGCTGCTCAGCCGCTCCAGGTTGGAGCCACAGACCCAGCCGACCCGGCT 1920
Db 1861 CTTCTCTGCTGCTGCTCAGCCGCTCCAGGTTGGAGCCACAGACCCAGCCGACCCGGCT 1920
QY 1921 GGGTCTGCGCTCTTCTGCTGCTTCCCTCCAGAAATGCGCTCAGACCTAGAAAGCTCA 1980
Db 1921 GGGTCTGCGCTCTTCTGCTGCTTCCCTCCAGAAATGCGCTCAGACCTAGAAAGCTCA 1980
QY 1981 ACCCCCTATGAGGCGCAGCTCTGGGGTAGCTCTGACCTCGACCTTATGTCAAAT 2040
Db 1981 ACCCCCTATGAGGCGCAGCTCTGGGGTAGCTCTGACCTCGACCTTATGTCAAAT 2040
QY 2041 TCACACCCATGTTTTTCAATTTGACCCGCGCTCTGCTCATATGACACCCAGCTCC 2100
Db 2041 TCACACCCATGTTTTTCAATTTGACCCGCGCTCTGCTCATATGACACCCAGCTCC 2100
QY 2101 TTTGAGAGGATCAGAGCCCATTTGCCAAGAGAGCGCTGCCAACATCTTGTCTCG 2160
Db 2101 TTTGAGAGGATCAGAGCCCATTTGCCAAGAGAGCGCTGCCAACATCTTGTCTCG 2160
QY 2161 ATTGCAAAATGACACCCAGTAATCTAGAACTTCTCAAGCCCTTTAACTCAGATGCA 2220
Db 2161 ATTGCAAAATGACACCCAGTAATCTAGAACTTCTCAAGCCCTTTAACTCAGATGCA 2220
QY 2221 AGCCACCGGGCAACCCCGCTCAATACCTCCCAAGAGAGGATGAGATATGTGACCTCACT 2280
Db 2221 AGCCACCGGGCAACCCCGCTCAATACCTCCCAAGAGAGGATGAGATATGTGACCTCACT 2280
QY 2281 GCTCCCGGCAACCCCGCTCAATACCTCCCAAGAGAGGATGAGATATGTGACCTCACT 2340
Db 2281 GCTCCCGGCAACCCCGCTCAATACCTCCCAAGAGAGGATGAGATATGTGACCTCACT 2340
QY 2341 AGGCTCAGAAATGAGCTCTGAAATAGGAGAGCTTAGCAAGAGAGAGATACAGGGTATCGG 2400
Db 2341 AGGCTCAGAAATGAGCTCTGAAATAGGAGAGCTTAGCAAGAGAGAGATACAGGGTATCGG 2400
QY 2401 GGGTTGAGTGTTCAGAACTCATTCGGGAAGATTAATCCAGTGTGCGCGCAGCCAC 2460
Db 2401 GGGTTGAGTGTTCAGAACTCATTCGGGAAGATTAATCCAGTGTGCGCGCAGCCAC 2460
QY 2461 CTGCAATCAAGCTTGGACAGCGGTTCTGTTCCGGAGGCAAAATTTCCCTAGGAAAA 2520
Db 2461 CTGCAATCAAGCTTGGACAGCGGTTCTGTTCCGGAGGCAAAATTTCCCTAGGAAAA 2520
QY 2521 GAAGACAGACTTTTCTAATGTGTCCAAATGCGGATCACTGTGAGATGGAATCTAGAG 2580
Db 2521 GAAGACAGACTTTTCTAATGTGTCCAAATGCGGATCACTGTGAGATGGAATCTAGAG 2580
QY 2581 CACTGAGCTCCCTGTCTGGAAGTATTAGAAAAAGCTGGCCAGGCAAGATGGCTCA 2640
Db 2581 CACTGAGCTCCCTGTCTGGAAGTATTAGAAAAAGCTGGCCAGGCAAGATGGCTCA 2640
QY 2641 GCGCTGTAATCCAGACTTTGGAGGCGGAGCGAGATCACTCTCTAAATAACAAAAATAGCA 2700
Db 2641 GCGCTGTAATCCAGACTTTGGAGGCGGAGCGAGATCACTCTCTAAATAACAAAAATAGCA 2700
QY 2701 GAGAACAGCTGGCCCAACATGTGTGAACCTCTCTCTAAATAACAAAAATAGCA 2760
Db 2701 GAGAACAGCTGGCCCAACATGTGTGAACCTCTCTCTAAATAACAAAAATAGCA 2760
QY 2761 GCGCTGAGGAGGAGGCTGCTGTAATCCAGTACTTTGGAGGCTGAGGATGAGATCACT 2820
Db 2761 GCGCTGAGGAGGAGGCTGCTGTAATCCAGTACTTTGGAGGCTGAGGATGAGATCACT 2820
```

```
QY 2821 TAAACCTGAGGCGAGAGGTTACAGTGGCCAGAGATCGTGCCACTGCTTCAGGCTGG 2880
Db 2821 TAAACCTGAGGCGAGAGGTTACAGTGGCCAGAGATCGTGCCACTGCTTCAGGCTGG 2880
QY 2881 CGACAGAGCAAGACTCTCTCTCAAAAAA 2918
Db 2881 CGACAGAGCAAGACTCTCTCTCAAAAAA 2918

RESULT 2
ADRI6281
ID ADRI6281 standard; cDNA; 2918 BP.
XX AC ADRI6281;
XX DT 21-OCT-2004 (first entry)
XX DE Human Cayman ataxia variant cDNA.
XX KW Human; jittery; ss; gene; Cayman ataxia; ATCAY; KIAA1872;
XX KW Chromosome 19p13.3; ataxia; myoclonus; dystonia; epilepsy; myasthenia;
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX CDS 373..1488
XX FT /*tag= a
XX FT /product= "Cayman ataxia protein"
XX FT replace(1275,C)
XX FT /*tag= b
XX FT /standard_name= "single nucleotide polymorphism"
XX PN US2004146900-A1.
XX PD 29-JUL-2004.
XX PF 03-NOV-2003; 2003US-00699941.
XX PR 01-NOV-2002; 2002US-0422971P.
XX PR 08-NOV-2002; 2002US-0424973P.
XX PA (UNMI ) UNIV MICHIGAN.
XX PI Burmeister M;
XX WPI; 2004-552665/53.
XX P-PSDB; ADRI6282.
XX Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in
XX subject, by detecting presence or absence of variant Cayman ataxia
XX polypeptide or nucleic acid in biological sample.
XX Claim 5; SEQ ID NO 8; 98pp; English.
XX The invention relates to detecting variant Cayman ataxia polypeptide or
XX nucleic acid sequence (ATCAY) in a subject, involving providing a
XX biological sample from a subject, where the biological sample comprises a
XX Cayman ataxia polypeptide or nucleic acid, and detecting the presence or
XX absence of a variant Cayman ataxia polypeptide or nucleic acid in the
XX biological sample. Also included is a kit comprising a reagent for
XX detecting the presence or absence of a variant Cayman ataxia nucleic acid
XX or polypeptide in a biological sample. The human ATCAY gene (previously
XX isolated as cDNA KIAA1872, located on chromosome 19p13.3) was identified
XX and mapped using the fact that it is the homologue of the mouse jittery
XX gene, located on mouse chromosome 10. The method is useful for detecting
XX variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.
XX The presence of the variant Cayman ataxia polypeptide or nucleic acid is
XX indicative of a disorder chosen from ataxia, myoclonus, dystonia,
XX epilepsy, and myasthenia in the subject. The biological sample is chosen
XX from blood sample, a tissue sample, urine sample, saliva sample, and an
XX amniotic fluid sample. The subject is chosen from embryo, foetus, newborn
```

CC animal, young animal, and an adult animal. The animal is a human. The
CC human is an adult female of child-bearing age. The present sequence is
CC the cDNA encoding the variant Human ATCAY protein, the allele encoding
CC the disease causing protein.
XX
SQ Sequence 2918 BP; 699 A; 867 C; 798 G; 554 T; 0 U; 0 Other;

Query Match 99.9%; Score 2916.4; DB 13; Length 2918;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2917; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGACGAGCTACCTCGAGGAGCGGCGCA 60
DB 1 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGACGAGCTACCTCGAGGAGCGGCGCA 60
QY 61 GCGCGGCGGCGATGAGAGGGGGGCGCAGCGCCCGCGCTGGGAGGCCACCGCTAAC 120
DB 61 GCGCGGCGGCGATGAGAGGGGGGCGCAGCGCCCGCGCTGGGAGGCCACCGCTAAC 120
QY 121 CTTGCAACCCCAACCCCTTGCAAAAGAGCTGGCGGGCTGGCCACGTCGCCCTGGG 180
DB 121 CTTGCAACCCCAACCCCTTGCAAAAGAGCTGGCGGGCTGGCCACGTCGCCCTGGG 180
QY 181 TGACCTTCTCGATCGAGATCGGCCCTCGAGCATCTCTTCTCTAGGCTCTGAA 240
DB 181 TGACCTTCTCGATCGAGATCGGCCCTCGAGCATCTCTTCTCTAGGCTCTGAA 240
QY 241 GCGCCGGGAGCGTGAGCGATGCCAGCTGCACCCGGGAGGGCTCGCTTTGTTGCCA 300
DB 241 GCGCCGGGAGCGTGAGCGATGCCAGCTGCACCCGGGAGGGCTCGCTTTGTTGCCA 300
QY 301 GTAAGGAGGAGGCTGTCTAGCTGCAAGGGGTTCATCCCTGCTTCAAGCCAGTGCCTC 360
DB 301 GTAAGGAGGAGGCTGTCTAGCTGCAAGGGGTTCATCCCTGCTTCAAGCCAGTGCCTC 360
QY 361 TTCCAGCTCCCATGGGACCAAGCAAGCCAGCTCGGATGGAAACGTGGACGTGAAG 420
DB 361 TTCCAGCTCCCATGGGACCAAGCAAGCCAGCTCGGATGGAAACGTGGACGTGAAG 420
QY 421 GAGGAATGGCAGGACGAAGATCTTCCAGGCCACTCCAGAGAGAGCGGGTGGAACTG 480
DB 421 GAGGAATGGCAGGACGAAGATCTTCCAGGCCACTCCAGAGAGAGCGGGTGGAACTG 480
QY 481 CTTGGCAGCCCGTGGAAAGACATCTCTCTCTCCCAACACGCTAAATTTCAACGGAGCG 540
DB 481 CTTGGCAGCCCGTGGAAAGACATCTCTCTCTCCCAACACGCTAAATTTCAACGGAGCG 540
QY 541 CATCGTAAGGAGGAGCGTGGTGGCCCGCAGAGATCAACATTTCTCGATCAGAGTGAG 600
DB 541 CATCGTAAGGAGGAGCGTGGTGGCCCGCAGAGATCAACATTTCTCGATCAGAGTGAG 600
QY 601 GGGTCCCTGCTGTCGATGACTTTTGGATACCCCTGATGACCTGGATATTACGTGGAT 660
DB 601 GGGTCCCTGCTGTCGATGACTTTTGGATACCCCTGATGACCTGGATATTACGTGGAT 660
QY 661 GACATCGAGACCCCGATGAGACCGACTCGTGGAGTTCTTGGGGAATGGCAACGACTG 720
DB 661 GACATCGAGACCCCGATGAGACCGACTCGTGGAGTTCTTGGGGAATGGCAACGACTG 720
QY 721 GAGTGGGAGACACACCCCGTGGCCACCGCCAGAAACATGCCCCGGGAGACGCGGAT 780
DB 721 GAGTGGGAGACACACCCCGTGGCCACCGCCAGAAACATGCCCCGGGAGACGCGGAT 780
QY 781 CTATTTGGGAGCGGACGAGGAGCGGAGCGCGCCGCAACGCGGCGCTGTGGCGGACA 840
DB 781 CTATTTGGGAGCGGACGAGGAGCGGAGCGCGCCGCAACGCGGCGCTGTGGCGGACA 840
QY 841 GTGATCATCGGGAGCAAGAGCACCGTATAGACTTGCAATGATCCGGCTTACATGAAA 900
DB 841 GTGATCATCGGGAGCAAGAGCACCGTATAGACTTGCAATGATCCGGCTTACATGAAA 900
QY 901 GTGGTCAACCCAGGGGTACTAGGGGAGGGCTCAACGCCCATCATGCTCTTCGAGCC 960
DB 901 GTGGTCAACCCAGGGGTACTAGGGGAGGGCTCAACGCCCATCATGCTCTTCGAGCC 960

DB 901 GTGGTCAACCCAGGGGTACTAGGGGAGGGCTCAACGCCCATCATGCTCTTCGAGCC 960
QY 961 TGCTTCTCTTCCAGACAGCAGCGCTCCCGACTACCACTACATCATCGAGAACCTCTTCTCG 1020
DB 961 TGCTTCTCTTCCAGACAGCAGCGCTCCCGACTACCACTACATCATCGAGAACCTCTTCTCG 1020
QY 1021 TACGTCAATCAGCAGCTTAGAGCTCTGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1080
DB 1021 TACGTCAATCAGCAGCTTAGAGCTCTGTGGCTGAGGACTACATGATCGTGTACCTGAAC 1080
QY 1081 GGTGCCACGCCCCCGCGGAGGATGCTGGAATCGGCTGGCTGAAGAGTGTACCAAGATG 1140
DB 1081 GGTGCCACGCCCCCGCGGAGGATGCTGGAATCGGCTGGCTGAAGAGTGTACCAAGATG 1140
QY 1141 ATCGACCGGAGGTTCGGGAAAAAAGCTGAAGTCTTTGATCATCGTCCACCTCTGGTTC 1200
DB 1141 ATCGACCGGAGGTTCGGGAAAAAAGCTGAAGTCTTTGATCATCGTCCACCTCTGGTTC 1200
QY 1201 ATTGGGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAAGATC 1260
DB 1201 ATTGGGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTTCATCAACAAGATC 1260
QY 1261 CAGTACGTGCAAGCTTGGAGACCTCGAGGACATCATCTCCCTATGGAACACGTCCAGATC 1320
DB 1261 CAGTACGTGCAAGCTTGGAGACCTCGAGGACATCATCTCCCTATGGAACACGTCCAGATC 1320
QY 1321 CCAGACTGCTGCTGCAATACGAGAGGAAAGACTGAAGGCGAGGAGGAGCGCGAGG 1380
DB 1321 CCAGACTGCTGCTGCAATACGAGAGGAAAGACTGAAGGCGAGGAGGAGCGCGAGG 1380
QY 1381 CCCCAGCCGAGTTTGTGCTGCCAGGTCTGAAAGAAAGCCAGAGGTGGCACAGTGGAA 1440
DB 1381 CCCCAGCCGAGTTTGTGCTGCCAGGTCTGAAAGAAAGCCAGAGGTGGCACAGTGGAA 1440
QY 1441 AACAGGTCTGCTGTGGTCTCAGAGATCAGGAAACAGCATGTCTGAGGCGACGTGAGC 1500
DB 1441 AACAGGTCTGCTGTGGTCTCAGAGATCAGGAAACAGCATGTCTGAGGCGACGTGAGC 1500
QY 1501 ATAAACAAAGGACATGAAAGAGATTCAGATGSCCAAGAAACCTCTGTGAGCGCCACTG 1560
DB 1501 ATAAACAAAGGACATGAAAGAGATTCAGATGSCCAAGAAACCTCTGTGAGCGCCACTG 1560
QY 1561 GCGCCAGATCTCATCTGCTCATCTGAGTCCCAATCTTCCAGGGGTGCGAGCCCTCC 1620
DB 1561 GCGCCAGATCTCATCTGCTCATCTGAGTCCCAATCTTCCAGGGGTGCGAGCCCTCC 1620
QY 1621 GTTCACTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAAGAACATGTGATTTTTTTTTT 1680
DB 1621 GTTCACTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAAGAACATGTGATTTTTTTTTT 1680
QY 1681 TAAACGATCAGTATTTGTGCGTTCCAGAAAAAGGCGCCAGCTCTGAGCCCTCACCTTCC 1740
DB 1681 TAAACGATCAGTATTTGTGCGTTCCAGAAAAAGGCGCCAGCTCTGAGCCCTCACCTTCC 1740
QY 1741 ACATCTACGAACTCTCAGCCGAGGAAAGCAAGAGCGAGGGGTGGCCCGCGCTG 1800
DB 1741 ACATCTACGAACTCTCAGCCGAGGAAAGCAAGAGCGAGGGGTGGCCCGCGCTG 1800
QY 1801 GGTGGCCTCCGCTCTGCTCGAGCCCTGTGCTGAGAGCTGGATACAAGATTCAAGACC 1860
DB 1801 GGTGGCCTCCGCTCTGCTCGAGCCCTGTGCTGAGAGCTGGATACAAGATTCAAGACC 1860
QY 1861 CTTCTCTTGTGTCACCCGCTTCCAGGTTGGAGCCACAGACCCACCGCCACCCCGGT 1920
DB 1861 CTTCTCTTGTGTCACCCGCTTCCAGGTTGGAGCCACAGACCCACCGCCACCCCGGT 1920
QY 1921 GGGTCTGCTCTTTCTGTGCTTTCCCTCCAGAAATCGGCTCAGACCTAGAAAGTCA 1980
DB 1921 GGGTCTGCTCTTTCTGTGCTTTCCCTCCAGAAATCGGCTCAGACCTAGAAAGTCA 1980
QY 1981 ACCCCCCCTATGAGGGCGACGCTCTGGGGTGTGCTCTGAGCTCCGACCTTATGTCCAAAT 2040
DB 1981 ACCCCCCCTATGAGGGCGACGCTCTGGGGTGTGCTCTGAGCTCCGACCTTATGTCCAAAT 2040

QY 2041 TCACACCCATGTTTTCATTGACCGCGCCCTTCTGCTCATATGACACCCAGCTCC 2100
DB |||||
QY 2041 TCACACCCATGTTTTCATTGACCGCGCCCTTCTGCTCATATGACACCCAGCTCC 2100
DB |||||
QY 2101 TTTGAGAGGATCAGAGCCCATTCACAGAGAGCGCTGCCAACCATCTTGTCTCTCG 2160
DB |||||
QY 2101 TTTGAGAGGATCAGAGCCCATTCACAGAGAGCGCTGCCAACCATCTTGTCTCTCG 2160
DB |||||
QY 2161 ATTGCAAAATGACACCCAGTAACTCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTC 2220
DB |||||
QY 2161 ATTGCAAAATGACACCCAGTAACTCTAGAACATTTCTCAAGCCCTTTAACTCAGATGTC 2220
DB |||||
QY 2221 AGCCACCGGGCAAAACCCGTTCAATACCTCCCAAGGAATGAGATATGAGACCTCACT 2280
DB |||||
QY 2221 AGCCACCGGGCAAAACCCGTTCAATACCTCCCAAGGAATGAGATATGAGACCTCACT 2280
DB |||||
QY 2281 GCTCCCGCCCAACCCAGGCTGAGGACAGCGCAAGCGCTGTTCCGGGTGGACAGCAG 2340
DB |||||
QY 2281 GCTCCCGCCCAACCCAGGCTGAGGACAGCGCAAGCGCTGTTCCGGGTGGACAGCAG 2340
DB |||||
QY 2341 AGGCTCAGAACTGGCTCTGAAATAGGCGAGACCTAGCAAGAGAGATACAGGTTATCG 2400
DB |||||
QY 2341 AGGCTCAGAACTGGCTCTGAAATAGGCGAGACCTAGCAAGAGAGATACAGGTTATCG 2400
DB |||||
QY 2401 GCGTTTGAAGTCTCAGAACTGTCAGGAGATTAATCCAGTGCCTGCGCGCAGCCAC 2460
DB |||||
QY 2401 GCGTTTGAAGTCTCAGAACTGTCAGGAGATTAATCCAGTGCCTGCGCGCAGCCAC 2460
DB |||||
QY 2461 CTGCATTCAAGCTTGGACAGCGGCTTCTGTTGGGAGGCAATTTCCCTAGGAAAA 2520
DB |||||
QY 2461 CTGCATTCAAGCTTGGACAGCGGCTTCTGTTGGGAGGCAATTTCCCTAGGAAAA 2520
DB |||||
QY 2521 GAAGCAGACTTTTCTAAATGTGTCACAAATGCGATCACTGTGTCAGATGACTCTAGAG 2580
DB |||||
QY 2521 GAAGCAGACTTTTCTAAATGTGTCACAAATGCGATCACTGTGTCAGATGACTCTAGAG 2580
DB |||||
QY 2581 CACTGAGCTCCCTGCTCTGGAAGTATTTAAGAAAGGCTGGGCGCAGGACCATGGCTCA 2640
DB |||||
QY 2581 CACTGAGCTCCCTGCTCTGGAAGTATTTAAGAAAGGCTGGGCGCAGGACCATGGCTCA 2640
DB |||||
QY 2641 CGCCTGTAATCCAGACTTTGGAGCGCCAGGCGCGGATCACTGAGGTGAGGATTT 2700
DB |||||
QY 2641 CGCCTGTAATCCAGACTTTGGAGCGCCAGGCGCGGATCACTGAGGTGAGGATTT 2700
DB |||||
QY 2701 GAGAACAGCTGCGCAACATGTGAACTCTATCTACTTAAATATACAAAATTAGCCA 2760
DB |||||
QY 2701 GAGAACAGCTGCGCAACATGTGAACTCTATCTACTTAAATATACAAAATTAGCCA 2760
DB |||||
QY 2761 GCGGTGTTGCGAGGCTGCTGTAATCCAGTACTTGGAGGCTGAGGCAATGAGATCACT 2820
DB |||||
QY 2761 GCGGTGTTGCGAGGCTGCTGTAATCCAGTACTTGGAGGCTGAGGCAATGAGATCACT 2820
DB |||||
QY 2821 TAAACTGAGAGCGAGAGGTTACAGTGAGCCAGATCGGCCACTGCAATTCAGGCTGG 2880
DB |||||
QY 2821 TAAACTGAGAGCGAGAGGTTACAGTGAGCCAGATCGGCCACTGCAATTCAGGCTGG 2880
DB |||||
QY 2881 CCACAGAGCAGACTCTGTCTCAAAAAA 2918
DB |||||
QY 2881 CCACAGAGCAGACTCTGTCTCAAAAAA 2918
DB |||||

RESULT 3

ADQ85643

ID ADQ85643 standard; cDNA; 4978 BP.

XX AC ADQ85643;

XX DT 07-OCT-2004 (first entry)

DE Human tumour-associated antigenic target (TAT) cDNA sequence #2457.

XX KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

KW cancer; cell proliferative disorder; gene; ss.
OS Homo sapiens.
XX WO2004060270-A2.
XX 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
XX (GETH) GENENTECH INC.
XX (WUTD/) WU T D.
XX (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
XX WPI; 2004-534300/51.
XX New nucleic acid molecule and encoded polypeptide, for diagnosing, preventing or treating cell proliferative disorders such as cancer.
XX Claim 1; SEQ ID NO 2457; 5504pp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cDNA sequence from the present invention.

Sequence 4978 BP; 1241 A; 1416 C; 1366 G; 955 T; 0 U; 0 Other;

Query Match 99.9%; Score 2913.8; DB 12; Length 4978;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2915; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGACAGCTACTCTCGAGGCGGCGGCCA 60

DB 5 GCCGAGCCTCTGCCAGCCCTGAGCTGGGAAGACAGCTACTCTCGAGGCGGCGGCCA 64

Db 425 GAGAAATGGCAGGACGAAGATCTTTCCAGGCCACTCCCAAGAGAGCGGGGTGGAACTG 484
Qy 481 CTTGGCAGCCGGTGGGAAGACACATCTCTCTCTCCCAACACACGCTAAATTTCAACGAGCG 540
Db 485 CTTGGCAGCCGGTGGGAAGACACATCTCTCTCTCCCAACACACGCTAAATTTCAACGAGCG 544
Qy 541 CATCGTAAGAGGAGACGCTGGTGGGCCCAAGAGATCAACATTTCTCTGATCAGAGTGAG 600
Db 545 CATCGTAAGAGGAGACGCTGGTGGGCCCAAGAGATCAACATTTCTCTGATCAGAGTGAG 604
Qy 601 GGGTCCCTGCTGTCGATGATCTTTGGATACCCCTGATGACCTGATATTAACGTGGAT 660
Db 605 GGGTCCCTGCTGTCGATGATCTTTGGATACCCCTGATGACCTGATATTAACGTGGAT 664
Qy 661 GACATCGAGACCCCGATGAGACCGGACTCGCTGGAGTTCTTGGGGAAATGGCAACGAATG 720
Db 665 GACATCGAGACCCCGATGAGACCGGACTCGCTGGAGTTCTTGGGGAAATGGCAACGAATG 724
Qy 721 GAGTGGGAAGACGACACCCCGTGGGCAACCGCCCAAGAAACATGCCCCGGGACGCGCGAT 780
Db 725 GAGTGGGAAGACGACACCCCGTGGGCAACCGCCCAAGAAACATGCCCCGGGACGCGCGAT 784
Qy 781 CTATTTGGGACGGCAGCGAGGAGCGGACGCGCCCAAGAAACATGCCCCGGGACGCGCGAT 840
Db 785 CTATTTGGGACGGCAGCGAGGAGCGGACGCGCCCAAGAAACATGCCCCGGGACGCGCGAT 844
Qy 841 GTGATCATCGGAGACGAAGACACCGTATAGACCTGCACATGATCGGGCTTACATGAAG 900
Db 845 GTGATCATCGGAGACGAAGACACCGTATAGACCTGCACATGATCGGGCTTACATGAAG 904
Qy 901 GTGGTCAACCAACGGAAGGTAACGCGGAAGGCTCAACGCCCATCATCTGTTTCGAGCC 960
Db 905 GTGGTCAACCAACGGAAGGTAACGCGGAAGGCTCAACGCCCATCATCTGTTTCGAGCC 964
Qy 961 TGGTCTCTTCCAGACGAGCCTCCCGACTACCACTACATCATGAGAAACCTCTTCCTG 1020
Db 965 TGGTCTCTTCCAGACGAGCCTCCCGACTACCACTACATCATGAGAAACCTCTTCCTG 1024
Qy 1021 TAGCTCATCAGCAGCTTAGAGCTCTGTTGGCTGAGGACTACATGATCGTGACCTGAAC 1080
Db 1025 TAGCTCATCAGCAGCTTAGAGCTCTGTTGGCTGAGGACTACATGATCGTGACCTGAAC 1084
Qy 1081 GGTGCCACGCCCGCGGAGGATGCTTGGAAATCGGCTGAGGAGTGTCAACAGATG 1140
Db 1085 GGTGCCACGCCCGCGGAGGATGCTTGGAAATCGGCTGAGGAGTGTCAACAGATG 1144
Qy 1141 ATCGACCGGAGGTTGGGAAACCTTGAAGTCTTGTATCATGTCACCCCTCGTGGTTC 1200
Db 1145 ATCGACCGGAGGTTGGGAAACCTTGAAGTCTTGTATCATGTCACCCCTCGTGGTTC 1204
Qy 1201 ATTCCGACTGTGTGGCCATCTCTGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATC 1260
Db 1205 ATTCCGACTGTGTGGCCATCTCTGCCCTTTTCATCAGCGTCAAGTTTCATCAACAGATC 1264
Qy 1261 CAGTACGTGCACAGCTTGGAAACCTTGGAGCACTCATCCCTATGAAACACGCTCCAGATC 1320
Db 1265 CAGTACGTGCACAGCTTGGAAACCTTGGAGCACTCATCCCTATGAAACACGCTCCAGATC 1324
Qy 1321 CCAGACTGGCTCTGCATACGAAGAGGAAGTCAAGGCCGAGGAGGAGAGCGGAGG 1380
Db 1325 CCAGACTGGCTCTGCATACGAAGAGGAAGTCAAGGCCGAGGAGGAGAGCGGAGG 1384
Qy 1381 CCCACGCGGAGTTGTGTGCCAGTCTGAAGAGAGCCAGAGGTGGCACACAGTGGAA 1440
Db 1385 CCCACGCGGAGTTGTGTGCCAGTCTGAAGAGAGCCAGAGGTGGCACACAGTGGAA 1444
Qy 1441 AACAGTCTGCTGTGCTCAGAGATCAGGAACCAAGCATGTCTTGGGCGGACGTGAGC 1500
Db 1445 AACAGTCTGCTGTGCTCAGAGATCAGGAACCAAGCATGTCTTGGGCGGACGTGAGC 1504
Qy 1501 ATAAACAGGACATGGAAGAGATTCAGATGCGCAAGAACCTCTGTGAGCGCCACTG 1560

Db 1505 ATAAACAGGACATGGAAGAGATTCAGATGCGCAAGAAACCTCTGTGAGACGCCACTG 1564
Qy 1561 GCCCCAGATCTCATCTGCTCATCTGAGTCCCATACTTCCAAAGGTTGCCAGCCCTCC 1620
Db 1565 GCCCCAGATCTCATCTGCTCATCTGAGTCCCATACTTCCAAAGGTTGCCAGCCCTCC 1624
Qy 1621 GTTCAATCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAAGAAATTTGATTTTTTTTTT 1680
Db 1625 GTTCAATCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAAGAAATTTGATTTTTTTTTT 1684
Qy 1681 TAAACATGAGTATTTTGTGCGTTCCAGAAAGGGGCCAGCTCTGAGCCCTCACCTTCC 1740
Db 1685 TAAACATGAGTATTTTGTGCGTTCCAGAAAGGGGCCAGCTCTGAGCCCTCACCTTCC 1744
Qy 1741 ACACTCAGCAACTCTCAGCCGAGGAAGCAAGAACGCGAGGGGTGGCCCGGTGGCGTC 1800
Db 1745 ACACTCAGCAACTCTCAGCCGAGGAAGCAAGAACGCGAGGGGTGGCCCGGTGGCGTC 1804
Qy 1801 GGTGCTCTCCGCTCTCTGCTCGAGCCCTGTGTGTCAGAGCTGGATACAAGATTTCAAGACC 1860
Db 1805 GGTGCTCTCCGCTCTCTGCTCGAGCCCTGTGTGTCAGAGCTGGATACAAGATTTCAAGACC 1864
Qy 1861 CTTCTCTTGTCTGTCAACCGCTCAGGTTGAGCCACAGACCCACCGCCACCCCGGT 1920
Db 1865 CTTCTCTTGTCTGTCAACCGCTCAGGTTGAGCCACAGACCCACCGCCACCCCGGT 1924
Qy 1921 GGGTCTGCGTCTTTCTGCTGCTTCCCTCCAGAAATCGGCTCAGACCTTAGAAGCTCA 1980
Db 1925 GGGTCTGCGTCTTTCTGCTGCTTCCCTCCAGAAATCGGCTCAGACCTTAGAAGCTCA 1984
Qy 1981 ACCCCCTTATGAGGGCCACGCTCTGGGTAGTCTCTGACCTCCGACCTTATGTCCAAAT 2040
Db 1985 ACCCCCTTATGAGGGCCACGCTCTGGGTAGTCTCTGACCTCCGACCTTATGTCCAAAT 2044
Qy 2041 TCACACCATGTTTTCATTTGACCCCGCTCTCTGCTCATATGACACCCAGCTCC 2100
Db 2045 TCACACCATGTTTTCATTTGACCCCGCTCTCTGCTCATATGACACCCAGCTCC 2104
Qy 2101 TTTGAGAGGATCAGAGCCCATTTGCAAGAAAGAGCGCTGCCCAACCATCTTGTCTCCG 2160
Db 2105 TTTGAGAGGATCAGAGCCCATTTGCAAGAAAGAGCGCTGCCCAACCATCTTGTCTCCG 2164
Qy 2161 ATTGCAAAATGACACCCAGTAATCTAGAACATTTCAAGCCCTTTAACTCAGATGTCA 2220
Db 2165 ATTGCAAAATGACACCCAGTAATCTAGAACATTTCAAGCCCTTTAACTCAGATGTCA 2224
Qy 2221 AGCCACCGGGCAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGACCTCACT 2280
Db 2225 AGCCACCGGGCAACCCCGTCAATACCTCCACCAAGGAATGAGATATGTGACCTCACT 2284
Qy 2281 GCTCCCCCAACCCAGCGTCAGGCTGGGACACGCCAACCTGTTCCGGGTGGAAACAGAG 2340
Db 2285 GCTCCCCCAACCCAGCGTCAGGCTGGGACACGCCAACCTGTTCCGGGTGGAAACAGAG 2344
Qy 2341 AGGCTCAGAAATCTGCTCTGAAATAGGCAAGCACTAGCAAGAGGAAGATACAGGGTATCGG 2400
Db 2345 AGGCTCAGAAATCTGCTCTGAAATAGGCAAGCACTAGCAAGAGGAAGATACAGGGTATCGG 2404
Qy 2401 GGGTTTGAAGTTTTCAGAAAGTCAATTCGGGAAGATAAATCCAGTGGCGTGGCCGACGAC 2460
Db 2405 GGGTTTGAAGTTTTCAGAAAGTCAATTCGGGAAGATAAATCCAGTGGCGTGGCCGACGAC 2464
Qy 2461 CTGCATTCAAAGCTTGGACACGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA 2520
Db 2465 CTGCATTCAAAGCTTGGACACGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA 2524
Qy 2521 GAAGACAGATTTTCTAAATGTGGTCCAAATGCGGATCACTGTCAGATGGACTCTAGAAG 2580
Db 2525 GAAGACAGATTTTCTAAATGTGGTCCAAATGCGGATCACTGTCAGATGGACTCTAGAAG 2584
Qy 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACGATGGCTCA 2640
Db 2585 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCCAGGACGATGGCTCA 2644

QY 1021 TACGTATCAGCAGCTTAGAGCTCCTGGTGGCTGAGGACTATATGATCGTGTACCTGAAC 1080
DB 1025 TACGTATCAGCAGCTTAGAGCTCCTGGTGGCTGAGGACTATATGATCGTGTACCTGAAC 1084
QY 1081 GGTGCCACGCCCGCGGAGGATCGCTGGAATCGGCTGGCTGAAGAAGTGTACACAGATG 1140
DB 1085 GGTGCCACGCCCGCGGAGGATCGCTGGAATCGGCTGGCTGAAGAAGTGTACACAGATG 1144
QY 1141 ATCGACCGGAGGTTCGGGAAAACTCGAAGTCTCTTGATCATCGTCCACCCCTCGTGGTTC 1200
DB 1145 ATCGACCGGAGGTTCGGGAAAACTCGAAGTCTCTTGATCATCGTCCACCCCTCGTGGTTC 1204
QY 1201 ATTGCGACTGTGTGCGCATCTCTGCGCTTTTCATCAGCGTCAAGTTTCATCAACAAGATC 1260
DB 1205 ATTGCGACTGTGTGCGCATCTCTGCGCTTTTCATCAGCGTCAAGTTTCATCAACAAGATC 1264
QY 1261 CAGTACGTGCACAGCTTGGGAAGACCTGGAGCACTCATCCCTATGGAACACGCTCCAGATC 1320
DB 1265 CAGTACGTGCACAGCTTGGGAAGACCTGGAGCACTCATCCCTATGGAACACGCTCCAGATC 1324
QY 1321 CCAGACTGCGTCTCTGCAATACGAAGAGGAAGACTGAAGGCCAGAGGAGGAGCGGAGG 1380
DB 1325 CCAGACTGCGTCTCTGCAATACGAAGAGGAAGACTGAAGGCCAGAGGAGGAGCGGAGG 1384
QY 1381 CCCAGCGGAGTTTGTGTCGCCAGGTCTGAAGAGAGCAGAGGTGGCAACAGTGGAA 1440
DB 1385 CCCAGCGGAGTTTGTGTCGCCAGGTCTGAAGAGAGCAGAGGTGGCACCAGTGGAA 1444
QY 1441 AACAGTCTGCTCTGCTCAGAGATCAGGAACCAAGCATGTCCTGAGCGGACGTGAGC 1500
DB 1445 AACAGTCTGCTCTGCTCAGAGATCAGGAACCAAGCATGTCCTGAGCGGACGTGAGC 1504
QY 1501 ATAAAGAGGACATGAAGAAAGATTCAGATGCCAGAAACCTCTGTCAGAGCCCACTG 1560
DB 1505 ATAAAGAGGACATGAAGAAAGATTCAGATGCCAGAAACCTCTGTCAGAGCCCACTG 1564
QY 1561 GCCCAGATCTCATCTGCTGCTCATCTGAGTCCCAATCTTTCAGAGGTGCCAGCCCTCC 1620
DB 1565 GCCCAGATCTCATCTGCTGCTCATCTGAGTCCCAATCTTTCAGAGGTGCCAGCCCTCC 1624
QY 1621 GTTCATCTGAAACCCAGCATCTTTTCAGCTGCTTGAACCATTTGATTTT 1680
DB 1625 GTTCATCTGAAACCCAGCATCTTTTCAGCTGCTTGAACCATTTGATTTT 1682
QY 1681 TAAAGATGCAATTTTGTGGTTCAGAAAGGGCCAGCTCTGAGCCCTCAACCTTCC 1740
DB 1683 TAAAGATGCAATTTTGTGGTTCAGAAAGGGCCAGCTCTGAGCCCTCAACCTTCC 1742
QY 1741 AACTCAGCACTCTCAGCGGAGGAAGGCAAGGCGCAGGGGTGGCCGGTGGCGTC 1800
DB 1743 AACTCAGCACTCTCAGCGGAGGAAGGCAAGGCGCAGGGGTGGCCGGTGGCGTC 1802
QY 1801 GGTGGCTCCGCTCTGCTGCGAGCCCTCTGCTGAGCTGGATCAAGATTCAAGACC 1860
DB 1803 GGTGGCTCCGCTCTGCTGCGAGCCCTCTGCTGAGCTGGATCAAGATTCAAGACC 1862
QY 1861 CTTCCTCTGCTGCTCACCCTCCAGGTGGAGCCACAGACACCCCGCCCGGCT 1920
DB 1863 CTTCCTCTGCTGCTCACCCTCCAGGTGGAGCCACAGACACCCCGCCCGGCT 1922
QY 1921 GGTGCTGCTGCTCTTCTGCTGCTTCCCTCAGAAATGGGCTCAGACTAGAGCTCA 1980
DB 1923 GGTGCTGCTGCTCTTCTGCTGCTTCCCTCAGAAATGGGCTCAGACTAGAGCTCA 1982
QY 1981 ACCCCCTATGAGGGCCAGCTCTGGGTAGCTCTGAGCTCGGACCTTATGTCGAAT 2040
DB 1983 ACCCCCTATGAGGGCCAGCTCTGGGTAGCTCTGAGCTCGGACCTTATGTCGAAT 2042
QY 2041 TCACACCCATGGTTTTCATTTGACCCGCCCCCTTCTGCTCATATGACACCCAGCTCC 2100
DB 2043 TCACACCCATGGTTTTCATTTGACCCGCCCCCTTCTGCTCATATGACACCCAGCTCC 2102
QY 2101 TTTGAGAGGATCAGAGCCCATTTGCAACAAGAGCGCTGCGCAACCATCTTGTCTCCG 2160

DB 2103 TTTGAGAGGATCAGAGCCCATTTGCAACAAGAGCGCTGCGCAACCATCTTGTCTCCG 2162
QY 2161 ATTGCAAAATGACACCCAGTAATCTAGAACATTTCAAGCCCTTTAACTCAGATGTCA 2220
DB 2163 ATTGCAAAATGACACCCAGTAATCTAGAACATTTCAAGCCCTTTAACTCAGATGTCA 2222
QY 2221 AGCCACCCGGGAAAAACCCGCTCAATACCTCCCAACCAAGGAATGAGATATGTGACCTCACT 2280
DB 2223 AGCCACCCGGGAAAAACCCGCTCAATACCTCCCAACCAAGGAATGAGATATGTGACCTCACT 2282
QY 2281 GCTCCCCCAACCCAGCGTCAAGCTGGGACACGCCAACGCTGTTCCGGTGTGAACAGCAG 2340
DB 2283 GCTCCCCCAACCCAGCGTCAAGCTGGGACATGCCAACGCTGTTCCGGTGTGAACAGCAG 2342
QY 2341 AGGCTCAGAACTGGCTCTGAAATAGGACCTTAGCAAGAGGAAGATACAGGGTATCGG 2400
DB 2343 AGGCTCAGAACTGGCTCTGAAATAGGACCTTAGCAAGAGGAAGATACAGGGTATCGG 2402
QY 2401 GGGTTTGAAGTGTTCAGAAAGTCATTCGGGAAGATAAATCCAGTGGCTGGCGCAGCCAC 2460
DB 2403 GGGTTTGAAGTGTTCAGAAAGTCATTCGGGAAGATAAATCCAGTGGCTGGCGCAGCCAC 2462
QY 2461 CTGCATTCAAAGCTTTGGACCAAGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA 2520
DB 2463 CTGCATTCAAAGCTTTGGACCAAGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA 2522
QY 2521 GAAGACAGACTTTTCTAATGTGGTCCAAATCGGATCACTGGTCAGATGGACTCTAGAAG 2580
DB 2523 GAAGACAGACTTTTCTAATGTGGTCCAAATCGGATCACTGGTCAGATGGACTCTAGAAG 2582
QY 2581 CACTCAGCTCCCTGCTCTCTGGAAGTATTTTAAAGAAAGGCTGGGCGCAGGCACGATGGCTCA 2640
DB 2583 CACTCAGCTCCCTGCTCTCTGGAAGTATTTTAAAGAAAGGCTGGGCGCAGGCACGATGGCTCA 2642
QY 2641 GCGCTGTATCCAGACTTTTGGAGGCGCAGGCGGATCACTGAGGTGAGGAGTTT 2700
DB 2643 GCGCTGTATCCAGACTTTTGGAGGCGCAGGCGGATCACCTGAGGTGAGGAGTTT 2702
QY 2701 GAGAACGCTGGCCCAACATGGTGAACCTCATCTACTATAAATAAATAAATAAATTAGCCA 2760
DB 2703 GAGAACGCTGGCCCAACATGGTGAACCTCATCTACTATAAATAAATAAATAAATTAGCCA 2762
QY 2761 GCGCTGTGGGAGGCGCTGTAAATCCAGCTACTTTGGAGGCTGAGGCATGAGAACTCACT 2820
DB 2763 GCGCTGTGGGAGGCGCTGTAAATCCAGCTACTTTGGAGGCTGAGGCATGAGAACTCACT 2822
QY 2821 TAAACCTGAGAGGCGAGAGGTTACAGTGAGCAGAAATCGTGCACTGCAATCCAGCCTGG 2880
DB 2823 TAAACCTGAGAGGCGAGAGGTTACAGTGAGCAGAAATCGTGCACTGCAATCCAGCCTGG 2882
QY 2881 CGACAGCGCAAGACTCTCTCNAAAAAA 2917
DB 2883 CGACAGCGCAAGACTCTCTCNAAAAAA 2919

RESULT 6
AAD00128
ID AAD00128 standard; cDNA; 3065 BP.
XX
AC AAD00128;
XX
DT 31-JUL-2000 (first entry)
XX
DE Human NIP2b cDNA.
XX
KW NIP2b; human; NIP2 family; apoptosis; adenoviral E1B 19 kDa protein;
Bcl-2 protein; cell survival; signal transduction pathway; modulator;
KW programmed cell death; screening; diagnosis; prognosis; therapeutic;
KW pharmacogenetic; prophylaxis; drug; treatment; 88.
OS Homo sapiens.
XX

FH	Key	Location/Qualifiers
FT	5'UTR	1..369
FT		/tag= a
FT	CDS	370..1485
FT		/tag= b
FT		/product= "Human NIP2b protein"
FT		/note= "This region is specifically claimed as
FT	3'UTR	Seq.ID.NO:3"
FT		1486..3065
FT		/tag= c
XX		
PN	WO200023470-A1.	
XX		
XX	27-APR-2000.	
DD		
PF	19-OCT-1999;	99WO-US024335.
PR	19-OCT-1998;	98US-00174937.
XX	(MILL-) MILLENNIUM PHARM INC.	
FA	Curtis RAJ, Glucksmann MA;	
PI		
XX		
DR	WPI; 2000-339653/29.	
DR	P-PSDB; AAY70847.	
PT	Nucleic acids encoding NIP2 proteins family members designated NIP2b,	
PT	NIP2cL and NIP2eS, useful for preventing, diagnosing and treating	
PT	disorders associated with inappropriate NIP2 expression.	
XX		
PS	Claim 1a; Fig 1; 120pp; English.	
XX		
CC	The present sequence is the cDNA encoding the human NIP2b protein, a	
CC	member of the NIP2 protein family. It is predominantly expressed in the	
CC	brain and also in prostate and small intestine. It plays a major role in	
CC	apoptosis. The NIP2b nucleotide sequence is 74% identical to human mRNA	
CC	for KIAA0367. NIP2b interacts with proteins like adenoviral E1B 19 kDa	
CC	protein or BCL-2 protein to provide a cell survival function. It	
CC	activates NIP2b-dependent signal transduction pathway and modulates	
CC	programmed cell death. This sequence can be used to screen for NIP2b,	
CC	NIP2cL and NIP2eS sequences and mutations in them, in diagnostic and	
CC	prognostic assays, pharmacogenetics and for therapeutic and prophylactic	
CC	purposes. It can also be used to screen for naturally occurring	
CC	substrates, drugs or modulators of their activity. The proteins can be	
CC	used to treat disorders associated with programmed cell death	
XX		
SQ	Sequence 3065 BP; 738 A; 906 C; 839 G; 582 T; 0 U; 0 Other;	
	Query Match	99.0%; Score 2887.6; DB 3; Length 3065;
	Best Local Similarity	99.9%; Pred. No. 0;
	Matches 2890; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
QY	24	CTGGGAAGACGACTACCTCGGAGGCAGGGCGCGCGCGCGCGCATGAGGGGGC 83
Db	21	CCGGAGAAGACGACTACCTCGGAGGCAGGGCGCGCGCGCGCGCATGAGGGGGC 80
QY	84	GCAAGCGCAGCCCCCGCGTGGGGAGGCCAACCCGTAAACCCTGCACCCCACCCATCCCTGCA 143
Db	81	GCAGCGCAGCCCCCGCGTGGGGAGGCCAACCCGTAAACCCTGCACCCCACCCATCCCTGCA 140
QY	144	CAAAAGAGCTGGCGGGCGTGCGCACGTCGCCCTGGGGTGAATTCTTCGAGTAGCAATC 203
Db	141	CAAAAGAGCTGGCGGGCGTGCGCACGTCGCCCTGGGGTGAATTCTTCGAGTAGCAATC 200
QY	204	CGCCCTTCGAGCATCTCTTCTTCCTAGGCTCTGAAGCCCGGGAGCGTGAAGGATGC 263
Db	201	CGCCCTTCGAGCATCTCTTCTTCCTAGGCTCTGAAGCCCGGGAGCGTGAAGGATGC 260
QY	264	CCAGCTGCACCCCGGCGAGGGCTCGCCCTTTGTTGTCAGTAAGGAGGAGGCTGTCTCAG 323
Db	261	CCAGCTGCACCCCGGCGAGGGCTCGCCCTTTGTTGTCAGTAAGGAGGAGGCTGTCTCAG 320
QY	324	CTCAGAGGGGTGATCCCTGCTTCAAGCCAGTGCCTCTTCCAGCTCCCTCCATGGGGACAC 383

Db 1401 CAGGCTCTGAGAGAGGACGACAGAGTGGGACCAAGTGGAAAAACAGGCTCTGCTCTCTGCTCTCAGA 1460
Qy 1464 AGATCAGGAAACAGACATGTCTCTGAGCGACGCTGAGCATATAACAAAGGACATGGAAGAGA 1523
Db 1461 AGATCAGGAAACAGACATGTCTCTGAGCGACGCTGAGCATATAACAAAGGACATGGAAGAGA 1520
Qy 1524 TTCCAGATGCCAGAAACCTCTCTGACAGCCCACTGGGCCCAAGATCTCATCTGCTCTCA 1583
Db 1521 TTCCAGATGCCAGAAACCTCTCTGACAGCCCACTGGGCCCAAGATCTCATCTGCTCTCA 1580
Qy 1584 TCTGAGTCCCAATCTTCCAAAGGTCGCCGCCCTCCGTTCTCTCTGAAACCCAGCATC 1643
Db 1581 TCTGAGTCCCAATCTTCCAAAGGTCGCCGCCCTCCGTTCTCTCTGAAACCCAGCATC 1640
Qy 1644 CTTTTCAGCTGCTTGAACATTTGATTTTTTTTTTTTAAACGATGAGTATTTGTCGTT 1703
Db 1641 CTTTTCAGCTGCTTGAACATTTGATTTTTTTTTTTTAAACGATGAGTATTTGTCGTT 1700
Qy 1704 CAGAAAGGCGCCAGCTCTGAGCCCTCAACCTTCCACACTCAAGAACTCTCAGCCGAG 1763
Db 1701 CAGAAAGGCGCCAGCTCTGAGCCCTCAACCTTCCACACTCAAGAACTCTCAGCCGAG 1760
Qy 1764 GAAGGCAAGAGCGCAGGGGTGGCCGGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 1823
Db 1761 GAAGGCAAGAGCGCAGGGGTGGCCGGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 1820
Qy 1824 GCCCTCTGTGTGACAGCTGATACAAAGATTCAGAGCCCTTCTCTGCTTTGTCAACCGCTC 1883
Db 1821 GCCCTCTGTGTGACAGCTGATACAAAGATTCAGAGCCCTTCTCTGCTTTGTCAACCGCTC 1880
Qy 1884 CAGGTTGAGCCACAGACACCCACCCCGCTGGGTCTGGCTCTCTCTGCTCTCTGCTCTG 1943
Db 1881 CAGGTTGAGCCACAGACACCCACCCCGCTGGGTCTGGCTCTCTCTGCTCTCTGCTCTG 1940
Qy 1944 TTTCCCTCAGAGTCGGGCTCAGACCTAGAGCTCAACCCCGCTATGAGGCGCAGCTCC 2003
Db 1941 TTTCCCTCAGAGTCGGGCTCAGACCTAGAGCTCAACCCCGCTATGAGGCGCAGCTCC 2000
Qy 2004 TGGGCTAGCTCTCTGACCTCCGACCTTATGTCCAAATTTTCAACCCCAATGTTTTTCAATTG 2063
Db 2001 TGGGCTAGCTCTCTGACCTCCGACCTTATGTCCAAATTTTCAACCCCAATGTTTTTCAATTG 2060
Qy 2064 ACCCGCCCTTCTCGCTCATATGACACCCAGCTCTCTTGTGAGAGGATCAGAGCCCATTTG 2123
Db 2061 ACCCGCCCTTCTCGCTCATATGACACCCAGCTCTCTTGTGAGAGGATCAGAGCCCATTTG 2120
Qy 2124 CACAGAGAGCGCTGCGCAACCATCTTGTCTCTCGATGCAAAATGACACCCCAAGTAA 2183
Db 2121 CACAGAGAGCGCTGCGCAACCATCTTGTCTCTCGATGCAAAATGACACCCCAAGTAA 2180
Qy 2184 TCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGTCAAGCCACCGGGCAAAACCCGTCAA 2243
Db 2181 TCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGTCAAGCCACCGGGCAAAACCCGTCAA 2240
Qy 2244 TACCTCCCAACCAAGGAATGAGATATGTGACCTCACTGTCTCCCCCAACCCAGCGTCAGGC 2303
Db 2241 TACCTCCCAACCAAGGAATGAGATATGTGACCTCACTGTCTCCCCCAACCCAGCGTCAGGC 2300
Qy 2304 TGGGACACCCACGCTGTTCCGGTTGGACAGCAGAGGCTCAGAACTGCTCTGAAA 2363
Db 2301 TGGGACACCCACGCTGTTCCGGTTGGACAGCAGCAGCAGGCTCAGAACTGCTCTGAAA 2360
Qy 2364 TAGGCAGACTAGCAAGAGAGATACAGGATATCGGGGTTTGTAGTGTTCAGAGTCA 2423
Db 2361 TAGGCAGACTAGCAAGAGAGATACAGGATATCGGGGTTTGTAGTGTTCAGAGTCA 2420
Qy 2424 TTCCGGAAGATAAATCCAGTGGCTGGCCGACGCACTGTCAATTCAAAGCTTTGGACCAAGC 2483
Db 2421 TTCCGGAAGATAAATCCAGTGGCTGGCCGACGCACTGTCAATTCAAAGCTTTGGACCAAGC 2480
Qy 2484 GGGTTCTCTGTCGGGAGGCAATTTCCCTAGGAAAGAGCAGACTTTTCTAATGCG 2543
Db 2481 GGGTTCTCTGTCGGGAGGCAATTTCCCTAGGAAAGAGCAGACTTTTCTAATGCGG 2540

Qy 2544 TCCAAATCGGATCACTGCTCAGATGGACTCTAGAAGCACTCAGCTCCCTGTCTCTGGAA 2603
Db 2541 TCCAAATCGGATCACTGCTCAGATGGACTCTAGAAGCACTCAGCTCCCTGTCTCTGGAA 2600
Qy 2604 GTATTTAAGAAAAGGCTGGCCAGGACGATGGCTCACGCTCTGTAATCCAGACTTTGGG 2663
Db 2601 GTATTTAAGAAAAGGCTGGCCAGGACGATGGCTCACGCTCTGTAATCCAGACTTTGGG 2660
Qy 2664 AGGCCGAGCGAGCGGATCACTGAGGTGAGGATTTGAGAACACGCTGGCCCAATGCT 2723
Db 2661 AGGCCGAGCGAGCGGATCACTGAGGTGAGGATTTGAGAACACGCTGGCCCAATGCT 2720
Qy 2724 GAAACCTCATCTCTACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2783
Db 2721 GAAACCTCATCTCTACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2780
Qy 2784 TCCAGCTACTTTGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGAGTTTAC 2843
Db 2781 TCCAGCTACTTTGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGAGTTTAC 2840
Qy 2844 AGTGAGCCAAAGATCGTGCCACTGCAATTCAGGCTGGGCGACAGACCAAGACTCTGTCTCA 2903
Db 2841 AGTGAGCCAAAGATCGTGCCACTGCAATTCAGGCTGGGCGACAGACCAAGACTCTGTCTCA 2900
Qy 2904 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2917
Db 2901 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2914

RESULT 7

AAH16751

ID AAH16751 standard; cDNA; 2894 BP.

XX AAH16751;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:15950.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 95JP-00248036.

XX 27-AUG-1999; 95JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00193767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-

XX length cDNAs defined in the specification, and for the detection and/or

XX diagnosis of the abnormality of the proteins encoded by the full-length

XX cDNAs.

XX Claim 8; SEQ ID NO 15950; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-

XX length cDNAs defined in the specification. Where a primer set comprises:

XX (a) an oligo-dT primer and an oligonucleotide complementary to the

XX

XX

1808 TCGCTCTCTGCTCGAGCCCTCTGTGTGCTCAGAGCTGGATCAAGATTCAAGACCTTCTCT 1867
1800 TCGCTCTCTGCTCGAGCCCTCTGTGTGCTCAGAGCTGGATCAAGATTCAAGACCTTCTCT 1859
1868 TCGCTCTCTGCTCGAGCCCTCTGTGTGCTCAGAGCTGGATCAAGATTCAAGACCTTCTCT 1927
1860 TCGCTCTCTGCTCGAGCCCTCTGTGTGCTCAGAGCTGGATCAAGATTCAAGACCTTCTCT 1919
1928 CGTCTCTCTGCTCGAGCCCTCTGTGTGCTCAGAGCTGGATCAAGATTCAAGACCTTCTCT 1987
1920 CGTCTCTCTGCTCGAGCCCTCTGTGTGCTCAGAGCTGGATCAAGATTCAAGACCTTCTCT 1979
1988 TATGAGGCGCAGCTCTGAGGCTAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCT 2047
1980 TATGAGGCGCAGCTCTGAGGCTAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCTGAGCTCT 2039
2048 CATGTTTTTCTATTTGAGCCGCTCTGCTGATATGATGATGATGATGATGATGATGATGATGAT 2107
2040 CATGTTTTTCTATTTGAGCCGCTCTGCTGATATGATGATGATGATGATGATGATGATGATGAT 2099
2108 GATCAGAGCCCATTCACAAAGAGAGCGCTGCCAACCATCTCTGCTCTCGATTGCAA 2167
2100 GATCAGAGCCCATTCACAAAGAGAGCGCTGCCAACCATCTCTGCTCTCGATTGCAA 2159
2168 ATGACACCCCGAGTATCTAGAACATCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2227
2160 ATGACACCCCGAGTATCTAGAACATCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2219
2228 GGGCAACCCCGTCAATCTCCACCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2287
2220 GGGCAACCCCGTCAATCTCCACCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2279
2288 CAACCCAGCGCTCAGGCTGGGACACAGCGCTGTTCCGGGTTGGAACAGAGAGCTCA 2347
2280 CAACCCAGCGCTCAGGCTGGGACACAGCGCTGTTCCGGGTTGGAACAGAGAGCTCA 2339
2348 GAACTGGCTCTGAATAGGACAGCTTAGCAAGAGAGATACAGGTTATCGGGGTTG 2407
2340 GAACTGGCTCTGAATAGGACAGCTTAGCAAGAGAGATACAGGTTATCGGGGTTG 2399
2408 AGTGTCTCAGAGTCAATTCGGGAGATTAATCAGTGCGCTGGCCGACGACCTGCAAT 2467
2400 AGTGTCTCAGAGTCAATTCGGGAGATTAATCAGTGCGCTGGCCGACGACCTGCAAT 2459
2468 CAAAGCTTGGACCCAGCGGTTCTTGTTCGGGAGCAATTTCCCTAGGAAAGAGAC 2527
2460 CAAAGCTTGGACCCAGCGGTTCTTGTTCGGGAGCAATTTCCCTAGGAAAGAGAC 2519
2528 GACTTTTCTAATGTGTCCAAATGCGGATCAGTGGTCAAGTGGATCTTAGAAGCACTGAG 2587
2520 GACTTTTCTAATGTGTCCAAATGCGGATCAGTGGTCAAGTGGATCTTAGAAGCACTGAG 2579
2588 CTCCTGTCTGGAAGTATTTAAGAAAGGCTGGCCAGGCGACGATGCTCAGCCTGT 2647
2580 CTCCTGTCTGGAAGTATTTAAGAAAGGCTGGCCAGGCGACGATGCTCAGCCTGT 2639
2648 AATCCAGACTTTGGGAGGCGGAGGAGCGGATCAGCTGAGTGAGGATTTTGAACA 2707
2640 AATCCAGACTTTGGGAGGCGGAGGAGCGGATCAGCTGAGTGAGGATTTTGAACA 2699
2708 GCTTGCCCAACATGTTGAAACCTCATCTCTACTTAAATAATACAAAATATGACGAGCGTGG 2767
2700 GCTTGCCCAACATGTTGAAACCTCATCTCTACTTAAATAATACAAAATATGACGAGCGTGG 2759
2768 TGGCAGGTGCTGTATCCAGCTACTTGGGAGGCTGAGGATGAGATCACTTAAACT 2827
2760 TGGCAGGTGCTGTATCCAGCTACTTGGGAGGCTGAGGATGAGATCACTTAAACT 2819
2828 GAGAGCAGAGGTTACAGTGAGCAAGATCGTGCCACTGCTATTCAGCTGCGGACACA 2887
2820 GAGAGCAGAGGTTACAGTGAGCAAGATCGTGCCACTGCTATTCAGCTGCGGACACA 2879
2888 GCAAGACTCTGTCTC 2902

Db 2880 GCAAGACTCTGTCTC 2894

RESULT 8

ADB62843

ID ADB62843 standard; cDNA; 2759 BP.

XX AC ADB62843;

DT 04-DEC-2003 (first entry)

XX Human cDNA encoding clone OCBBF20111370.

Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
tissue regeneration; cell regeneration; membrane protein;
signal transduction-related protein; transcription-related protein;
osteoporosis; neurological disease; cancer; tumour.

OS Homo sapiens.

XX Key Location/Qualifiers
CDS 296..1345

FT /*tag= a

XX EP1308459-A2.

XX 07-MAY-2003.

XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

XX 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehli S;

XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;

XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-450961/43.

XX P-PSDB; ADB64813.

XX New polynucleotides and polypeptides, useful for developing a diagnostic

XX marker or medicines for regulation of their expression and activity, or

XX as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected

XX from 1970 fully defined nucleotide sequences which encode novel

XX polypeptides. Also claimed is a polypeptide encoded by the polynucleotide

XX or its partial peptide, an antibody binding to the polypeptide or peptide

XX of the polynucleotide, immunologically assaying the polypeptide or

XX peptide of the polynucleotide by contacting the polypeptide or peptide

XX with the antibody of the encoded protein, and observing the binding

XX between the two, a transformant carrying the polynucleotide in an

XX expressible manner and an antisense polynucleotide. The oligonucleotide

XX is useful as a primer for synthesising the polynucleotide, or as a probe

XX for detecting the polynucleotide. The polynucleotides and encoded

XX proteins are useful as pharmaceutical agents and many disease-related

XX genes may be included in them, for developing a diagnostic marker or

XX medicines for regulation of their expression and activity, or as targets

XX of gene therapy. The genes are involved in tissue and/or cell

XX regeneration. Membrane proteins, signal transduction-related proteins,

XX transcription-related proteins, disease-related proteins and genes

XX encoding them can be used as indicators for diseases (e.g. osteoporosis,

XX neurological diseases, cancer, tumours). The cDNA may be used to regulate

XX the activity or expression of the encoded protein to treat diseases. The

XX sequence presented is a cDNA of the invention. Note: Some of the sequence

XX data for this patent is not represented in the printed specification, but

CC is based on sequence information supplied by the European Patent Office.

XX Sequence 2759 BP; 654 A; 821 C; 747 G; 537 T; 0 U; 0 Other;

SQ Query Match 89.0%; Score 2598.2; DB 10; Length 2759;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2756; Conservative 0; Mismatches 3; Indels 136; Gaps 2;

```
QY 8 CTCTGCAGCCCTGAGCTGGGAAGAAGACAGCTACTCTCGAGGCGAGGCGCGAGCGGGC 67
DB 1 CTCTGCAGCCCTGAGCTGGGAAGAAGACAGCTACTCTCGAGGCGAGGCGCGAGCGGGC 60
QY 68 GCGGATGAGAGGGGGCGAGCGCGAGCGCGCGCTGGGAGCCACCGCTAACCTGAC 127
DB 61 GCGGATGAGAGGGGGCGAGCGCGAGCGCGCGCTGGGAGCCACCGCTAACCTGAC 120
QY 128 CCACCCACCCCTGCACAAAGAGCTGGCGGGCGCTGGCCACGCTGCCCTGGGTGACCTT 187
DB 121 CCCACCCA----- 128
QY 188 CTTGGATGAGAATCCGCCCTTCGAGCATCTCTTCTCTAGGCTCTGAAGGCCCGG 247
DB 129 -----CCCTCGAGCATCTCTTCTCTAGGCTCTGAAGGCCCGG 170
QY 248 GAGCGTGAAGCATGCGGAGCTGCACCCGCGGCGAGGCTCGCTTTGTTGGCAGTAAGA 307
DB 171 GAGCGTGAAGCATGCGGAGCTGCACCCGCGGAGGCTCGCTTTGTTGGCAGTAAGA 230
QY 308 GAGAGGCTGTCTCAGCTGCAGAGGGGTCTATCCCTGCTTCAAGCCAGTCCCTTCCCG 367
DB 231 GAGAGGCTGTCTCAGCTGCAGAGGGGTCTATCCCTGCTTCAAGCCAGTCCCTTCCCG 290
QY 368 CTCCCATGGGACCAACCGAAGCCACGCTCCGATGGAACGCTGGAAGGAGGAAT 427
DB 291 CTCCCATGGGACCAACCGAAGCCACGCTCCGATGGAACGCTGGAAGGAGGAAT 350
QY 428 GGCAGGACGAAGATCTTCCAGGCCACTCCAGAGAGAGAGCGGGGTGGAACTGCTTGGCA 487
DB 351 GGC----- 353
QY 488 GCCCGGTGAAGACACATCTCTCTCCCAACAGCTAAATTTCAACGGAGCGCATCGTA 547
DB 354 -----AAGACATCTCTCTCTCCCAACAGCTAAATTTCAACGGAGCGCATCGTA 404
QY 548 AGAGGAAGACGCTGTGGTGGCCAGAGATCAACATTTCTCTGATCAGAGTGAGGGTCC 607
DB 405 AGAGGAAGACGCTGTGGTGGCCAGAGATCAACATTTCTCTGATCAGAGTGAGGGTCC 464
QY 608 TGCTGTCCGATGACTTTCTTGGATACCCCTGATGACTTAACTGATGATGATGATGATG 667
DB 465 TGCTGTCCGATGACTTTCTTGGATACCCCTGATGACTTAACTGATGATGATGATGATG 524
QY 668 AGACCCCGATCAGACCGGCTGCTGGATTTCTGGGATGCGCAACGATGAGTGGG 727
DB 525 AGACCCCGATCAGACCGGCTGCTGGATTTCTGGGATGCGCAACGATGAGTGGG 584
QY 728 AAGAGGACACCCCGTGGCCAGCGCAAGAACATGCGCGGGGACAGCGCGGATCTATTG 787
DB 585 AAGAGGACACCCCGTGGCCAGCGCAAGAACATGCGCGGGGACAGCGCGGATCTATTG 644
QY 788 GGGACGGCAGCAGGAGGAGCGGCGCGCAACCGGCGCTGTGGCGGACAGTGAATCA 847
DB 645 GGGACGGCAGCAGGAGGAGCGGCGCGCAACCGGCGCTGTGGCGGACAGTGAATCA 704
QY 848 TCGGGGAGCAAGAGCAGCTATAGACCTGCAATGATCGGCTTACATGAAGTGGTCA 907
DB 705 TCGGGGAGCAAGAGCAGCTATAGACCTGCAATGATCGGCTTACATGAAGTGGTCA 764
QY 908 CCCACGGAGGTAAGTACGGCGAAGGCTCAACGCCATCATCTCTTCCGAGCCTGCTTCC 967
DB 765 CCCACGGAGGTAAGTACGGCGAAGGCTCAACGCCATCATCTCTTCCGAGCCTGCTTCC 824
QY 968 TTCCAGACGAGCGCTCCCGGACTACCACTACATCATGGAGAACCTCTTCTGTGACGTCA 1027
```

```
DB 825 TTCCAGACGAGCGCTCCCGGACTACCACTACATCATGGAGAACCTCTTCTGTGACGTCA 884
QY 1028 TCAGCAGCTTTAGAGCTCTCTGTGGTGGAGGCTACATGATCGTGTACCTGGAAGGTCGCA 1087
DB 885 TCAGCAGCTTTAGAGCTCTCTGTGGTGGAGGCTACATGATCGTGTACCTGGAAGGTCGCA 944
QY 1088 CGCCCCGGCGGAGGATGCTTGGATCGGCTGCTGAAGAGTGTCTACCAAGATGATGACGAC 1147
DB 945 CGCCCCGGCGGAGGATGCTTGGATCGGCTGCTGAAGAGTGTCTACCAAGATGATGACGAC 1004
QY 1148 GAGGTTGCGGAAAAAAGCTTGAAGTCTTGTATCATCGTCCACCCCTGTGTGTTTCAATTCGA 1207
DB 1005 GAGGTTGCGGAAAAAAGCTTGAAGTCTTGTATCATCGTCCACCCCTGTGTGTTTCAATTCGA 1064
QY 1208 CTGTGCTGCGCATCTCTCGCCCTTTCATCAGGCTCAAGTTCATCAACAGATCCAGTAGG 1267
DB 1065 CTGTGCTGCGCATCTCTCGCCCTTTCATCAGGCTCAAGTTCATCAACAGATCCAGTAGG 1124
QY 1268 TGCAAGCTTTGGAAGACCTTGGAGCAACTCATCCCTATGGAACACAGTCCAGATCCAGACT 1327
DB 1125 TGCAAGCTTTGGAAGACCTTGGAGCAACTCATCCCTATGGAACACAGTCCAGATCCAGACT 1184
QY 1328 GCGTCTCTGCAATACGAAGAGAAAGACTGAAGGCGAGGAGGAGCGGAGGCCCCCAGC 1387
DB 1185 GCGTCTCTGCAATACGAAGAGAAAGACTGAAGGCGAGGAGGAGCGGAGGCCCCCAGC 1244
QY 1388 GCGAGTTTGTGCTGCCAGGCTTGAAGAGAACCCAGAGGTGGCACCAGTGGAAACAGGT 1447
DB 1245 GCGAGTTTGTGCTGCCAGGCTTGAAGAGAACCCAGAGGTGGCACCAGTGGAAACAGGT 1304
QY 1448 CTGCTCTGCTCTCAGAGATCAGGAAACAGCATGCTCTGAGGCGGAGCTGAGCATACAA 1507
DB 1305 CTGCTCTGCTCTCAGAGATCAGGAAACAGCATGCTCTGAGGCGGAGCTGAGCATACAA 1364
QY 1508 AGGACATGGAAGAGATTCAGATGCGCAGAAACCTCTGTGAGAGCGCCACCTGGGCCCGAG 1567
DB 1365 AGGACATGGAAGAGATTCAGATGCGCAGAAACCTCTGTGAGAGCGCCACCTGGGCCCGAG 1424
QY 1568 ATCTCATCTCTGCTCATCTGAGTCCCAATCTTCCAGGAGTGGCAGCGCCCTCGGTTCATC 1627
DB 1425 ATCTCATCTCTGCTCATCTGAGTCCCAATCTTCCAGGAGTGGCAGCGCCCTCGGTTCATC 1484
QY 1628 TCTGAAACCCAGCATCTCTTTTTCAGTCTGCTGAAACATTTGATTTTTTTTTTTTAAACAT 1687
DB 1485 TCTGAAACCCAGCATCTCTTTTTCAGTCTGCTGAAACATTTGATTTTTTTTTTTTAAACAT 1544
QY 1688 GCAGTATTTGTGCGTTCCAGAAAAAGGCGCCAGCTCTGAGCGCCCTCACCCCTTCCACATCA 1747
DB 1545 GCAGTATTTGTGCGTTCCAGAAAAAGGCGCCAGCTCTGAGCGCCCTCACCCCTTCCACATCA 1604
QY 1748 CGAATCTCAGCGGAGGAGCAAGAGCGAGGGGTGGCCCGCGTGGGTGGGTGGCC 1807
DB 1605 CGAATCTCAGCGGAGGAGCAAGAGCGAGGGGTGGCCCGCGTGGGTGGGTGGCC 1664
QY 1808 TCCGCTCTGCTCGCAGCGCCCTGTGTGCTCAGAGCTGATACAGATTTCAAGACCTTCTCT 1867
DB 1665 TCCGCTCTGCTCGCAGCGCCCTGTGTGCTCAGAGCTGATACAGATTTCAAGACCTTCTCT 1724
QY 1868 TGCTTGTCAACCGCTTCCAGGTTGGAGCCACAGACACCCAGCCCGCCCGCTGGGTCTG 1927
DB 1725 TGCTTGTCAACCGCTTCCAGGTTGGAGCCACAGACACCCAGCCCGCCCGCTGGGTCTG 1784
QY 1928 CGTCTTCTGTGCTCTTCCCTCCAGATCGGCGCTCAGACCTTAGAAGCTCAACCCCCC 1987
DB 1785 CGTCTTCTGTGCTCTTCCCTCCAGATCGGCGCTCAGACCTTAGAAGCTCAACCCCCC 1844
QY 1988 TATGAGGCGCAGCTCTCGGGGTAGCTCTGAGCTCTCGACCTTATGTTCAAAATTTTCAACC 2047
DB 1845 TATGAGGCGCAGCTCTCGGGGTAGCTCTGAGCTCTCGACCTTATGTTCAAAATTTTCAACC 1904
QY 2048 CATGGTTTTTTCATTTGACCCCGCCCTTCTCGCTCATATGACACCCAGCTCTTCTTGGAGA 2107
```

```
Db 1905 CATGGTTTTCATTTGACCGCGCCCTTCTCGCTCATATGACACCCAGCTCTTTTGAGA 1964
QY 2108 GGATCAGAGCCCATTTGCAAGAAGAGCCGCTCCCAACCATCTCTGCTCTCGATTGCAA 2167
Db 1965 GGATCAGAGCCCATTTGCAAGAAGAGCCGCTCCCAACCATCTCTGCTCTCGATTGCAA 2024
QY 2168 AATGACACCCAGTAACTTAGAATCTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2227
Db 2025 AATGACACCCAGTAACTTAGAATCTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2084
QY 2228 GGCAAAACCCGTCATCTCCCAAGGATGAGATATGTGGACCTCAGTCTCTCC 2287
Db 2085 GGCAAAACCCGTCATCTCCCAAGGATGAGATATGTGGACCTCAGTCTCTCC 2144
QY 2288 CAACCCAGGCTCAGGCTGGGACACGCGCAACGCTGTTCCGGTTGGAAACAGCAGAGCTCA 2347
Db 2145 CAACCCAGGCTCAGGCTGGGACATGCGCAACGCTGTTCCGGTTGGAAACAGCAGAGCTCA 2204
QY 2348 GAAACTGGCTCTGAAATAGGACAGCTTAGCAAGAGAAAGATACAGGGTATCGGGCGTTTG 2407
Db 2205 GAAACTGGCTCTGAAATAGGACAGCTTAGCAAGAGAAAGATACAGGGTATCGGGCGTTTG 2264
QY 2408 AGTGTCTCAGAGCTATCGGAGATTAATCCAGTCCGCTGGCCGACGACCTGCATT 2467
Db 2265 AGTGTCTCAGAGCTATCGGAGATTAATCCAGTCCGCTGGCCGACGACCTGCATT 2324
QY 2468 CAAAGCTTGACACGCGGTTCTTGTTCGGGAGGCAAAATTCCTTAGGAAAGAGACA 2527
Db 2325 CAAAGCTTGACACGCGGTTCTTGTTCGGGAGGCAAAATTCCTTAGGAAAGAGACA 2384
QY 2528 GACTTTTCTAATGTGTCAAATGCGGATCCTGTGTCAGATGACTCTAGAAACACTGAG 2587
Db 2385 GACTTTTCTAATGTGTCAAATGCGGATCCTGTGTCAGATGACTCTAGAAACACTGAG 2444
QY 2588 CTCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCGGACGATGCTCAGGCTGT 2647
Db 2445 CTCCTGTCTCTGGAAGTATTTAAGAAAGGCTGGGCGGACGATGCTCAGGCTGT 2504
QY 2648 AATCCAGACTTTGGGAGGCGGAGGCGGATCACCTGAGTGAGGATTTGAAACA 2707
Db 2505 AATCCAGACTTTGGGAGGCGGAGGCGGATCACCTGAGTGAGGATTTGAAACA 2564
QY 2708 GCTGCGCAACATGTTGAAACCTCATCTCTACTTAAATAATACAAATAATGACGAGCGTG 2767
Db 2565 GCTGCGCAACATGTTGAAACCTCATCTCTACTTAAATAATACAAATAATGACGAGCGTG 2624
QY 2768 TGGCAGGTGCTGTATCCAGTACTTGGAGGCTGAGGCGATGAGAACTTAAACCT 2827
Db 2625 TGGCAGGTGCTGTATCCAGTACTTGGAGGCTGAGGCGATGAGAACTTAAACCT 2684
QY 2828 GAGAGGACAGAGTTACAGTGAGCAAGATCGTGCCACTGCTATCCAGCTGGGCGACAGA 2887
Db 2685 GAGAGGACAGAGTTACAGTGAGCAAGATCGTGCCACTGCTATCCAGCTGGGCGACAGA 2744
QY 2888 GCAAGACTGTCTC 2902
Db 2745 GCAAGACTGTCTC 2759
```

RESULT 9
ADQ63969

ID ADQ63969 standard; cDNA; 2786 BP.

XX AC

XX ADQ63969;

XX XX

DT 07-OCT-2004 (first entry)

XX XX

DE Novel human cDNA sequence #1130.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.

XX Homo sapiens.

XX OS
XX EP1440981-A2.XX
XX 28-JUL-2004.

XX 21-JAN-2004; 2004BP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

XX 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX Yamamoto J, Isono Y, Nagai K, Irie R;

XX WPI; 2004-535376/52.

XX P-PSDB; ADQ66157.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,

XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 1130; 2449pp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded

XX polypeptides, sequences hybridizing to these nucleotides, sequences

XX encoding partial polypeptides and sequences having 70% or 90% identity to

XX the nucleotide and protein sequences. The nucleotides and polypeptides

XX are useful as diagnostic markers or therapeutic target for the diseases

XX or morbid states. They are also useful for treating osteoporosis,

XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,

XX dementia and various cancers. This sequence corresponds to a nucleotide

XX sequence of the invention.

XX SQ Sequence 2786 BP; 698 A; 795 C; 742 G; 551 T; 0 U; 0 Other;

XX Query Match 80.8%; Score 2357.6; DB 12; Length 2786;

XX Best Local Similarity 95.2%; Pred No. 0;

XX Matches 2492; Conservative 0; Mismatches 4; Indels 122; Gaps 1;

QY 300 AGTAAGGAGGAGGCTGTCTCAGCTGAGAGGGGTCTATCCCTGCTTCAAGCAGTGCT 359

Db 1 AGTAAGGAGGAGGCTGTCTCAGCTGAGAGGGGTCTATCCCTGCTTCAAGCAGTGCT 60

QY 360 CTTCCAGCTCCCATGGGACCAACCGAAGCCACGCTCCGATGGAACAGTGGACGTGA 419

Db 61 CTTCCAGCTCCCATGGGACCAACCGAAGCCACGCTCCGATGGAACAGTGGACGTGA 120

QY 420 GGAGGAATGGCAGGACGAAGATCTTCCAGGCGCACTCCCAAGAGAGCGGGGTGAACT 479

Db 121 GGAGGAATGGCAGGACGAAGATCTTCCAGGCGCACTCCCAAGAGAGCGGGGTGAACT 180

QY 480 GCTTGGCAGCGCGTGGGAAGACACATCTCTCTCCCAACAGCTTAAATTTCAACGGAGC 539

Db 181 GCTTGGCAGCGCGTGGGAAGACACATCTCTCTCCCAACAGCTTAAATTTCAACGGAGC 240

QY 540 GCATCGTAAGGAGGAGACGCTGGTGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGA 599

Db 241 GCATCGTAAGGAGGAGACGCTGGTGGCCCGCAGAGATCAACATTTCTCTGGATCAGAGTGA 300

QY 600 GGGGTCCCTGTCTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACTGGA 659

Db 301 GGGGTCCCTGTCTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACTGGA 338

QY 660 TGACATCGAGACCCCGGATGAGACCGACTCGTGGAGTCTCTGGGGAATGGCAACGAACT 719

Db 339 ----- 338

QY 720 GGAGTGGGAGAGACACCCCGTGGCCACCGCCCAAGAACATGCCCGGGGACAGCGCGGA 779

Db 339 -----ATGCCCGGGGACAGCGCGGA 358

QY 780 TCTATTGCGGACGCGACGAGGAGCGGACGCGCGCGCAACGCGGCGCTGTGGCGAC 839
DB 359 TCTATTGCGGACGCGACGAGGAGCGGACGCGCGCGCAACGCGGCGCTGTGGCGAC 418
QY 840 AGTGATCATCGGCGGACGAGGACGCGATAGACCTGCAATCATCGGCGCTTACATGAA 899
DB 419 AGTGATCATCGGCGGACGAGGACGCGATAGACCTGCAATCATCGGCGCTTACATGAA 478
QY 900 AGTGATCATCGGCGGACGAGGACGCGATAGACCTGCAATCATCGGCGCTTACATGAA 959
DB 479 AGTGATCATCGGCGGACGAGGACGCGATAGACCTGCAATCATCGGCGCTTACATGAA 538
QY 960 CTGCTTCTTCCAGACGAGGACGCGATAGACCTGCAATCATCGGCGCTTACATGAA 1019
DB 539 CTGCTTCTTCCAGACGAGGACGCGATAGACCTGCAATCATCGGCGCTTACATGAA 598
QY 1020 GTACGTCTCAGCAGCTTACAGCTCTCTGCTGAGGAGCTACATCATCGTGTACCTGAA 1079
DB 599 GTACGTCTCAGCAGCTTACAGCTCTCTGCTGAGGAGCTACATCATCGTGTACCTGAA 658
QY 1080 CGGTGCCACGCGCGCGGAGGAGCTTGAATCGGCTGCTGAGGAGCTTACATCATCGAT 1139
DB 659 CGGTGCCACGCGCGCGGAGGAGCTTGAATCGGCTGCTGAGGAGCTTACATCATCGAT 718
QY 1140 GATCGACCGGAGGTTGCGGAGAAACCTGAGTCTTGTGATCATCGTCCACCCCTGTGTT 1199
DB 719 GATCGACCGGAGGTTGCGGAGAAACCTGAGTCTTGTGATCATCGTCCACCCCTGTGTT 778
QY 1200 CATTCGGAGCTGTGCGGACCTCTCGCCCTTTCATCAGGCTCAAGTTCATCAACAGAT 1259
DB 779 CATTCGGAGCTGTGCGGACCTCTCGCCCTTTCATCAGGCTCAAGTTCATCAACAGAT 838
QY 1260 CCAGTACGTGCAAGCTTGGAGACCTGGAGCAACTCATCCCTATGGAACAGTTCAGAT 1319
DB 839 CCAGTACGTGCAAGCTTGGAGACCTGGAGCAACTCATCCCTATGGAACAGTTCAGAT 898
QY 1320 CCAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379
DB 899 CCAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958
QY 1380 GCGCCAGCGGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
DB 959 GCGCCAGCGGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1018
QY 1440 AAACAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1499
DB 1019 AAACAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
QY 1500 CATACAAAGGACATGGAAGAGATTCAGATGCGCAAGAAACCTCTGTGAGAGCGCCACT 1559
DB 1079 CATACAAAGGACATGGAAGAGATTCAGATGCGCAAGAAACCTCTGTGAGAGCGCCACT 1138
QY 1560 GCGCCAGATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619
DB 1139 GCGCCAGATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1198
QY 1620 CGTTCACTCTGAAACCCAGCATCTTTTTCAGTCTGCTGAAACATTTGTTTCTTTT 1679
DB 1199 CGTTCACTCTGAAACCCAGCATCTTTTTCAGTCTGCTGAAACATTTGTTTCTTTT 1258
QY 1680 TTAAAGATCAGTATTGTCGCTTCCAGAAAGGCGCCAGCTCTGAGCCCTTCAACCTTTC 1739
DB 1259 TTAAAGATCAGTATTGTCGCTTCCAGAAAGGCGCCAGCTCTGAGCCCTTCAACCTTTC 1318
QY 1740 CACACTCAGAACTCTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1799
DB 1319 CACACTCAGAACTCTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1378
QY 1800 CGGTGCTCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1859
DB 1379 CGGTGCTCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1438
QY 1860 CTTTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1919

RESULT 10

DB 1439 CCTTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1498
QY 1920 TGGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1979
DB 1499 TGGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1558
QY 1980 AACCCCTTATGAGGCGCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2039
DB 1559 AACCCCTTATGAGGCGCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1618
QY 2040 TTACACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2099
DB 1619 TTACACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1678
QY 2100 CTTTTCAGAGGATCAGAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2159
DB 1679 CTTTTCAGAGGATCAGAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1738
QY 2160 GATTGCAAAATGACACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219
DB 1739 GATTGCAAAATGACACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1798
QY 2220 AAGCCACCGGCGCAACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2279
DB 1799 AAGCCACCGGCGCAACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1858
QY 2280 TGCTCCCCCAACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2339
DB 1859 TGCTCCCCCAACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1918
QY 2340 GAGGCTCAGAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2399
DB 1919 GAGGCTCAGAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1978
QY 2400 GCGCTTTCAGTGTTCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2459
DB 1979 GCGCTTTCAGTGTTCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2038
QY 2460 CCGCTTTCAGTGTTCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2519
DB 2039 CCGCTTTCAGTGTTCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2098
QY 2520 AGAAGACAGACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2579
DB 2099 AGAAGACAGACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158
QY 2580 GCCTGAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2639
DB 2159 GCCTGAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2218
QY 2640 AGCTGTGTAATCCAGACTTTGAGGAGCGCGAGGCGGATCCTGAGGTGAGGAGTT 2699
DB 2219 AGCTGTGTAATCCAGACTTTGAGGAGCGCGAGGCGGATCCTGAGGTGAGGAGTT 2278
QY 2700 TGAGAACAGCTGCGCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2759
DB 2279 TGAGAACAGCTGCGCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2338
QY 2760 AGGCTGTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2819
DB 2339 AGGCTGTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2398
QY 2820 TTAAACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2879
DB 2399 TTAAACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2458
QY 2880 GCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2917
DB 2459 GCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2496

ADRL6283
ID ADR16283 standard; DNA; 47219 BP.
XX AC
XX ADR16283;
XX DT
XX 21-OCT-2004 (first entry)
XX DE Human Cayman ataxia splice site mutant DNA.
XX KW Human; jittery; ds; Cayman ataxia; ATCAY; KIAA1872; chromosome 19p13.3;
XX KM ataxia; myoclonus; dystonia; epilepsy; nyctagmus; splice site mutant;
XX KW SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT misc_feature 33181..33190
XX FT /*tag= a
XX FT /note= "Exon 9-intron 9 boundary"
XX FT replace(33187,G)
XX FT /*tag= b
XX FT /standard name= "Single nucleotide polymorphism"
XX FT /note= "Causes splice site mutation"
XX PN US2004146900-A1.
XX XX
XX PD 29-JUL-2004.
XX XX
XX PP 03-NOV-2003; 2003US-00699941.
XX XX
XX PR 01-NOV-2002; 2002US-0422971P.
XX PR 08-NOV-2002; 2002US-0424973P.
XX XX
XX PA (UNMI) UNIV MICHIGAN.
XX XX
XX PI Burmeister M;
XX XX
XX DR WPI; 2004-552665/53.
XX XX
XX PT Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in
XX PT subject by detecting presence or absence of variant Cayman ataxia
XX PT polypeptide or nucleic acid in biological sample.
XX XX
XX PS Claim 5; SEQ ID NO 10; 98pp; English.
XX XX
XX CC The invention relates to detecting variant Cayman ataxia polypeptide or
XX CC nucleic acid sequence (ATCAY) in a subject, involving providing a
XX CC biological sample from a subject, where the biological sample comprises a
XX CC Cayman ataxia polypeptide or nucleic acid, and detecting the presence or
XX CC absence of a variant Cayman ataxia polypeptide or nucleic acid in the
XX CC biological sample. Also included is a kit comprising a reagent for
XX CC detecting the presence or absence of a variant Cayman ataxia nucleic acid
XX CC or polypeptide in a biological sample. The human ATCAY gene (previously
XX CC isolated as cDNA KIAA1872, located on chromosome 19p13.3) was identified
XX CC and mapped using the fact that it is the homologue of the mouse jittery
XX CC gene, located on mouse chromosome 10. The method is useful for detecting
XX CC variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.
XX CC The presence of the variant Cayman ataxia polypeptide or nucleic acid is
XX CC indicative of a disorder chosen from ataxia, polypeptide or nucleic acid
XX CC epilepsy, and nyctagmus in the subject. The biological sample is chosen
XX CC from blood sample, a tissue sample, urine sample, saliva sample, and an
XX CC amniotic fluid sample. The subject is chosen from embryo, foetus, newborn
XX CC animal, young animal, and an adult animal. The animal is a human. The
XX CC human is an adult female of child-bearing age. The present sequence is
XX CC the DNA representing a Human ATCAY splice site mutant (disease causing)
XX CC where a single nucleotide polymorphism in the splice site of exon 9-
XX CC intron 9 causes a read through mutant protein to be produced.
XX SQ Sequence 47219 BP; 11562 A; 12203 C; 12045 G; 11209 T; 0 U; 200 Other;
Query Match 49.3%; Score 1438.2; DB 13; Length 47219;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1475 CAAGCATGTCTGTAGGCGACGTAGCATAAACAAGGACATGGAAGAGATTCAGATGCC 1534
DB |||||||
DB 43720 CTAGCATGTCTGTAGGCGACGTAGCATAAACAAGGACATGGAAGAGATTCAGATGCC 43779
QY 1535 AGAAAACTCTGTCTAGAGCCCACTGGCCCCAGATCTCATCTGCTCATCTCTGAGTCCC 1594
DB |||||||
DB 43780 AGAAAACTCTGTCTAGAGCCCACTGGCCCCAGATCTCATCTGCTCATCTCTGAGTCCC 43839
QY 1595 AATCTTCCAAAGGTCGCCAGCCCTCCGTTCTATCTCTGAAACCCAGCATCTCTTTCAGCTG 1654
DB |||||||
DB 43840 AATCTTCCAAAGGTCGCCAGCCCTCCGTTCTATCTCTGAAACCCAGCATCTCTTTCAGCTG 43899
QY 1655 CTTGAAAACTTGTATTTTTTTTTTAAACGATGAGTATTGTGCTTCCAGAAAAAGG 1714
DB |||||||
DB 43900 CTTGAAAACTTGTATTTTTTTTTTAAACGATGAGTATTGTGCTTCCAGAAAAAGG 43959
QY 1715 CCCAGCTCTGAGCCCTCTCACTCCCTTCCACACTCAAGAACTCTCAGCCGAGGAAGCAAGAA 1774
DB |||||||
DB 43960 CCCAGCTCTGAGCCCTCTCACTCCCTTCCACACTCAAGAACTCTCAGCCGAGGAAGCAAGAA 44019
QY 1775 GCGCAGGGGTGGCCCGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1834
DB |||||||
DB 44020 GCGCAGGGGTGGCCCGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 44079
QY 1835 CAGAGCTGGATACAAAGATTCAAGACCTTCTCTGCTTGTACCCCGCTCCAGGTTGGAGC 1894
DB |||||||
DB 44080 CAGAGCTGGATACAAAGATTCAAGACCTTCTCTGCTTGTACCCCGCTCCAGGTTGGAGC 44139
QY 1895 CACAGACACCCACCGCCACCCCGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1954
DB |||||||
DB 44140 CACAGACACCCACCGCCACCCCGGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 44199
QY 1955 AATGGCGCTCAGACTAGAGACTCAACCCCGCTATGAGGGCCAGCTCTGGGGTAGCTC 2014
DB |||||||
DB 44200 AATGGCGCTCAGACTAGAGACTCAACCCCGCTATGAGGGCCAGCTCTGGGGTAGCTC 44259
QY 2015 CTGACCTCGACCTTATGTCCAAATTTTCAACACCCATGGTTTTTTCATTTGACCGCCCT 2074
DB |||||||
DB 44260 CTGACCTCGACCTTATGTCCAAATTTTCAACACCCATGGTTTTTTCATTTGACCGCCCT 44319
QY 2075 TCTCGCTCATATATGACACCCAGCTCTTTGAGAGATCAGAGCCCATTTGCAACAAGAG 2134
DB |||||||
DB 44320 TCTCGCTCATATATGACACCCAGCTCTTTGAGAGATCAGAGCCCATTTGCAACAAGAG 44379
QY 2135 CCGCTGCCAACCATCTTGTCTCGATGCGAATGCAACCCAGTAACTTAGACATTT 2194
DB |||||||
DB 44380 CCGCTGCCAACCATCTTGTCTCGATGCGAATGCAACCCAGTAACTTAGACATTT 44439
QY 2195 CTCAAGCCCTTTTAACTCAGATGTCAAGCCACCGGGCAACCCCGTCAATACCTCCAC 2254
DB |||||||
DB 44440 CTCAAGCCCTTTTAACTCAGATGTCAAGCCACCGGGCAACCCCGTCAATACCTCCAC 44499
QY 2255 AAGGAATGAGATATGTGACCTCATCTGCTCCCAACCCAGCGTCAGGCTGGGACACGCC 2314
DB |||||||
DB 44500 AAGGAATGAGATATGTGACCTCATCTGCTCCCAACCCAGCGTCAGGCTGGGACACGCC 44559
QY 2315 AACGCTGTTCCGGGTGGNAACAGCAGAGGCTCAGAAATGCGCTCTGNAATAGGAGACCT 2374
DB |||||||
DB 44560 AACGCTGTTCCGGGTGGNAACAGCAGAGGCTCAGAAATGCGCTCTGNAATAGGAGACCT 44619
QY 2375 AGCAAGAGGAAGATACAGGGTATCGGGGCTTTGAGTGTTCAGAAAGTCAATTCGGGAAGAT 2434
DB |||||||
DB 44620 AGCAAGAGGAAGATACAGGGTATCGGGGCTTTGAGTGTTCAGAAAGTCAATTCGGGAAGAT 44679
QY 2435 AAATCCAGTGGCTGGCCGACCCACCTGCAATTTCAAAGCTTGGACAGCGGTTCTTGT 2494
DB |||||||
DB 44680 AAATCCAGTGGCTGGCCGACCCCTGCAATTTCAAAGCTTGGACAGCGGTTCTTGT 44739
QY 2495 CCGGAGGCGAATTTCCCTAGGAAAGAGACAGACTTTCTTAATGTGTGCCAATGGC 2554
DB |||||||
DB 44740 CCGGAGGCGAATTTCCCTAGGAAAGAGACAGACTTTCTTAATGTGTGCCAATGGC 44799

QY 2555 ATCACTGCTCAGATGGACTCTAGACGACCTGAGCTCCCTGTCTCTCGAAGTATTAAAGAA 2614
Db |||||
44800 ATCACTGCTCAGATGGACTCTAGACGACCTGAGCTCCCTGTCTCTCGAAGTATTAAAGAA 44859
QY 2615 AAGGCTGGGCGCAGCAGATGGCTCAGCGCTGTAATCCAGACTTTGGGAGGCCGAGGCA 2674
Db |||||
44860 AAGGCTGGGCGCAGCAGATGGCTCAGCGCTGTAATCCAGACTTTGGGAGGCCGAGGCA 44919
QY 2675 GGCGGATCACTCAGGTGAGGAGTTTGAGAACAGCCTGGCCCAACATGGTGAAACCTTCATC 2734
Db |||||
44920 GGCGGATCACTCAGGTGAGGAGTTTGAGAACAGCCTGGCCCAACATGGTGAAACCTTCATC 44979
QY 2735 TCTACTAAAATAACAAAATTAGCCAGGCTGTGTGAGGCTGCTTAAATCCAGTACT 2794
Db |||||
44980 TCTACTAAAATAACAAAATTAGCCAGGCTGTGTGAGGCTGCTTAAATCCAGTACT 45039
QY 2795 TGGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGCAGAGTTACAGTGAGCCAG 2854
Db |||||
45040 TGGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGCAGAGTTACAGTGAGCCAG 45099
QY 2855 ATCGTCCCACTGCATTCAGCTTGGGCGCAGAGCAAGACTCTGTCTCAAAAAAATAA 2914
Db |||||
45100 ATCGTCCCACTGCATTCAGCTTGGGCGCAGAGCAAGACTCTGTCTCAAAAAAATAA 45159
QY 2915 AAA 2917
Db |||||
45160 AAA 45162

RESULT 11
ADRI6284
ID ADR16284 standard; DNA; 59884 BP.
XX
AC ADR16284;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human Cayman ataxia genomic DNA.
XX
KW Human; jittery; ds; Cayman ataxia; ATCAY; KIAA1872; chromosome 19p13.3;
KW ataxia; myoclonus; dystonia; epilepsy; nystagmus; splice site mutant;
KW SNP; single nucleotide polymorphism; gene.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 45847..45852
FT /*tag= a
FT /note= "Exon 9-intron 9 boundary"
FT variation replace(45850,T)
FT /*tag= b
FT /standard_name= "single nucleotide polymorphism"
FT /note= "Causes splice site mutation"
XX
PN US2004146900-A1.
XX
PD 29-JUL-2004.
XX
PP 03-NOV-2003; 2003US-00699941.
XX
PR 01-NOV-2002; 2002US-0422971P.
XX
PR 08-NOV-2002; 2002US-0424973P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Burmeister M;
XX
DR WPI; 2004-552665/53.
XX
PT Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in
PT subject, by detecting presence or absence of variant Cayman ataxia
PT polypeptide or nucleic acid in biological sample.
XX

Claim 4; SEQ ID NO 11; 98pp; English.
The invention relates to detecting variant Cayman ataxia polypeptide or nucleic acid sequence (ATCAY) in a subject, involving providing a biological sample from a subject, where the biological sample comprises a Cayman ataxia polypeptide or nucleic acid, and detecting the presence or absence of a variant Cayman ataxia polypeptide or nucleic acid in the biological sample. Also included is a kit comprising a reagent for detecting the presence or absence of a variant Cayman ataxia nucleic acid or polypeptide in a biological sample. The human ATCAY gene (previously isolated as cDNA KIAA1872, located on chromosome 19p13.3) was identified and mapped using the fact that it is the homologue of the mouse Jittery gene, located on mouse chromosome 10. The method is useful for detecting variant Cayman ataxia polypeptide or nucleic acid sequence in a subject. The presence of the variant Cayman ataxia polypeptide or nucleic acid is indicative of a disorder chosen from ataxia, myoclonus, dystonia, epilepsy, and nystagmus in the subject. The biological sample is chosen from blood sample, a tissue sample, urine sample, saliva sample, and an amniotic fluid sample. The subject is chosen from embryo, foetus, newborn animal, young animal, and an adult animal. The animal is a human. The human is an adult female of child-bearing age. The present sequence is the human genomic DNA for ATCAY.

Sequence 59884 BP; 14656 A; 15794 C; 15192 G; 14042 T; 0 U; 200 Other;
Query Match 49.3%; Score 1438.2; DB 13; Length 59884;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1475 CAAGCATCTCTGAGGCGACGTGAGCATTAACAAGAGACATGAAGAATTCAGATGCC 1534
Db |||||
56385 CTAGCATCTCTGAGGCGACGTGAGCATTAACAAGAGACATGAAGAATTCAGATGCC 56444
QY 1535 AGAAAACCTCTGTCTGAGCGCCCATCTGCGCCAGATCTCATCTGCTCATCTGAGTCCC 1594
Db |||||
56445 AGAAAACCTCTGTCTGAGCGCCCATCTGCGCCAGATCTCATCTGCTCATCTGAGTCCC 56504
QY 1595 AATCTTCCAAGGTCAGCGCCCTCCGTTCTCTCTGAAACCCAGCATCTCTTCAGCTG 1654
Db |||||
56505 AATCTTCCAAGGTCAGCGCCCTCCGTTCTCTCTGAAACCCAGCATCTCTTCAGCTG 56564
QY 1655 CTGAAAACATTTGTAATTTTTTTTTTAAACGATGAGTATTGTCGTTCCAGAAAAGG 1714
Db |||||
56565 CTGAAAACATTTGTAATTTTTTTTTTAAACGATGAGTATTGTCGTTCCAGAAAAGG 56624
QY 1715 CCCAGCTCTGAGCCCTTCCACTTCCACTCTCAGCGAGGAGGCAAGAA 1774
Db |||||
56625 CCCAGCTCTGAGCCCTTCCACTTCCACTCTCAGCGAGGAGGCAAGAA 56684
QY 1775 GCGCAGGCGGTGGCGCGTGGTGGCTCCGCTCCGCTCCGCTCCGAGCCCTGTGGT 1834
Db |||||
56685 GCGCAGGCGGTGGCGCGTGGTGGCTCCGCTCCGCTCCGCTCCGAGCCCTGTGGT 56744
QY 1835 CAGAGCTGGATACAAGATTCAAGACCTTCTCTTGTGTCAACCGCTCCAGGTTGGAGC 1894
Db |||||
56745 CAGAGCTGGATACAAGATTCAAGACCTTCTCTTGTGTCAACCGCTCCAGGTTGGAGC 56804
QY 1895 CACAGACACCCACCGCCACCCCGGTGGTCTGGTCTCTTCTGTGCTTTCCTCCAG 1954
Db |||||
56805 CACAGACACCCACCGCCACCCCGGTGGTCTGGTCTCTTCTGTGCTTTCCTCCAG 56864
QY 1955 AATGGCGCTCAGACCTAGAGCTCAACCCCTATGAGGCCAGCTCTGGGTAGCTC 2014
Db |||||
56865 AATGGCGCTCAGACCTAGAGCTCAACCCCTATGAGGCCAGCTCTGGGTAGCTC 56924
QY 2015 CTGACCTCCGACCTTATGTCCAAATTTTCAACCCATGTTTTCATTTGACCCGCCCT 2074
Db |||||
56925 CTGACCTCCGACCTTATGTCCAAATTTTCAACCCATGTTTTCATTTGACCCGCCCT 56984
QY 2075 TCTCGCTCATATGACACCCAGCTCTTTTGAGAGGATCAGAGCCCATTTGCAAGAAG 2134
Db |||||
56985 TCTCGCTCATATGACACCCAGCTCTTTTGAGAGGATCAGAGCCCATTTGCAAGAAG 57044

```
QY 2135 CCGCTGCCACCATCTTGTCTCCGATTCGAAATGACACCCAGTAACTTAGAACATT 2194
|
|
|
Db 57045 CCGCTGCCACCATCTTGTCTCCGATTCGAAATGACACCCAGTAACTTAGAACATT 57104
|
|
|
QY 2195 CTCAGGCCCTTTAACTCAGATGTCAGGCGGCGGCAACCCGTCATACCTCCACC 2254
|
|
|
Db 57105 CTCAGGCCCTTTAACTCAGATGTCAGGCGGCGGCAACCCGTCATACCTCCACC 57164
|
|
|
QY 2255 AAGGAATGAGATATGTGGACCTCACTGCTCCCCCAACCCAGCGTCAGGCTGGGACACGCC 2314
|
|
|
Db 57165 AAGGAATGAGATATGTGGACCTCACTGCTCCCCCAACCCAGCGTCAGGCTGGGACACGCC 57224
|
|
|
QY 2315 AACGCTGTTCCGGGTTGGAACAGCAGAGGCTCAGAACTGGCTCTGAAATAGCAGACCT 2374
|
|
|
Db 57225 AACGCTGTTCCGGGTTGGAACAGCAGAGGCTCAGAACTGGCTCTGAAATAGCAGACCT 57284
|
|
|
QY 2375 AGCAAGAGGAAGATACAGGGTATCGGGCGGTTGAGTGTTCAGAGTCAATTCGGGAGAT 2434
|
|
|
Db 57285 AGCAAGAGGAAGATACAGGGTATCGGGCGGTTGAGTGTTCAGAGTCAATTCGGGAGAT 57344
|
|
|
QY 2435 AATCCAGTGCCTGCGCGGCGGACCTGCAATTCAGAGCTTGACACGCGGTTCTTGT 2494
|
|
|
Db 57345 AATCCAGTGCCTGCGCGGCGGACCTGCAATTCAGAGCTTGACACGCGGTTCTTGT 57404
|
|
|
QY 2495 CGGAGGCAAAATTTCCCTAGGAAAGAGACAGACTTTTCTAATGTGTCCAAATGCGG 2554
|
|
|
Db 57405 CGGAGGCAAAATTTCCCTAGGAAAGAGACAGACTTTTCTAATGTGTCCAAATGCGG 57464
|
|
|
QY 2555 ATCACTGGTCAGATGCACTTCAAGCACTGAGCTCCCTGCTCTGGAAGTATTTAGAA 2614
|
|
|
Db 57465 ATCACTGGTCAGATGCACTTCAAGCACTGAGCTCCCTGCTCTGGAAGTATTTAGAA 57524
|
|
|
QY 2615 AAGGCTGGGCGGACGATGCTCAGCGCTGTAATCCAGACTTTGGGAGGCCGAGGCA 2674
|
|
|
Db 57525 AAGGCTGGGCGGACGATGCTCAGCGCTGTAATCCAGACTTTGGGAGGCCGAGGCA 57584
|
|
|
QY 2675 GCGGATCACTGAGTGAGAGTTTGGAGAAAGAGCAAGCTTGCCCAACATGTTGAACTCATC 2734
|
|
|
Db 57585 GCGGATCACTGAGTGAGAGTTTGGAGAAAGAGCAAGCTTGCCCAACATGTTGAACTCATC 57644
|
|
|
QY 2735 TCTACTAAATAACAAATAATAGCCAGGCTGCTGAGGCTGCTGTAATCCGCTACT 2794
|
|
|
Db 57645 TCTACTAAATAACAAATAATAGCCAGGCTGCTGAGGCTGCTGTAATCCGCTACT 57704
|
|
|
QY 2795 TGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGAGTTACAGTGAGCCAG 2854
|
|
|
Db 57705 TGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGAGTTACAGTGAGCCAG 57764
|
|
|
QY 2855 ATGCTGCCACTGATTCAGGCTGGGCGGACAGAGCAAGACTCTGTCTCAAAAAAATAA 2914
|
|
|
Db 57765 ATGCTGCCACTGATTCAGGCTGGGCGGACAGAGCAAGACTCTGTCTCAAAAAAATAA 57824
|
|
|
QY 2915 AAA 2917
|
|
|
Db 57825 AAA 57827
```

RESULT 12

ADB62832

ID ADB62832 standard; cDNA; 2625 BP.

Query Match 45.4%; Score 1325.8; DB 10; Length 2625;
Best Local Similarity 99.8%; Pred. No. 3.8e-313;
Matches 1338; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

AC ADB62832;

XX 04-DEC-2003 (first entry)

QY 1562 CCCAGATCTCATCTGCTCATCTGAGTCCCAATCTTCCAGGGTCCAGCCCTCCG 1621

Db 1286 CTCAGATCTCATCTGCTCATCTGAGTCCCAATCTTCCAGGGTCCAGCCCTCCG 1345

QY 1622 TTCATCTCTGAAACCCAGCATCTTTTCAGTGTGTGAAACATGTATTTTTTTTTT 1681

Db 1346 TTCATCTCTGAAACCCAGCATCTTTTCAGTGTGTGAAACATGTATTTTTTTTTT 1404

QY 1682 AACGATGAGTATTTGCGGTTCAGAAAGGCGCCAGCTCTGAGCCCTCACCCTTCCA 1741

Db 1405 AACGATGAGTATTTGCGGTTCAGAAAGGCGCCAGCTCTGAGCCCTCACCCTTCCA 1464

OS Homo sapiens.

Key Location/Qualifiers
FH 1990. .2334
FT /*tag= a
FT /product= "Clone OCBBF20001780 protein"

PN EP1308459-A2.

XX 07-MAY-2003.

XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

PR 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

DR WPI; 2003-450961/43.
P-PSDB; ADB64802.

XX New polynucleotides and polypeptides, useful for developing a diagnostic
marker or medicines for regulation of their expression and activity, or
as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
from 1970 fully defined nucleotide sequences which encode novel
polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
or its partial peptide, an antibody binding to the polypeptide or peptide
of the polynucleotide, immunologically assaying the polypeptide or
peptide of the polynucleotide by contacting the polypeptide or peptide
with the antibody of the encoded protein, and observing the binding
between the two, a transformant carrying the polynucleotide in an
expressible manner and an antisense polynucleotide. The oligonucleotide
is useful as a primer for synthesizing the polynucleotide, or as a probe
for detecting the polynucleotide. The polynucleotides and encoded
proteins are useful as pharmaceutical agents and many disease-related
genes may be included in them, for developing a diagnostic marker or
medicines for regulation of their expression and activity, or as targets
of gene therapy. The genes are involved in tissue and/or cell
regeneration. Membrane proteins, signal transduction-related proteins,
transcription-related proteins, disease-related proteins and genes
encoding them can be used as indicators for diseases (e.g. osteoporosis,
neurological diseases, cancer, tumours). The cDNA may be used to regulate
the activity or expression of the encoded protein to treat diseases. The
sequence presented is a cDNA of the invention. Note: Some of the sequence
data for this patent is not represented in the printed specification, but
is based on sequence information supplied by the European Patent Office.

XX Sequence 2625 BP; 660 A; 754 C; 685 G; 526 T; 0 U; 0 Other;

1742 CACTCAGCACTCTCAGCCGAGGAGGCAAGAGCGGAGGGGGTGGCCCGCGTGGCGTGG 1801
1465 CACTCAGCACTCTCAGCCGAGGAGGCAAGAGCGGAGGGGGTGGCCCGCGTGGCGTGG 1524
1802 GTGGCCCTCCGCTCTCTGCTCCGAGCCCTCTGCTGGTGGTGGATCAAGATTCAGACCC 1861
1525 GTGGCCCTCCGCTCTCTGCTCCGAGCCCTCTGCTGGTGGTGGATCAAGATTCAGACCC 1584
1862 TTCTCTGCTGTGTCAACCGCTCCAGGTTGGAGCCACAGACACCCACCGCCACCCCGGCTG 1921
1585 TTCTCTGCTGTGTCAACCGCTCCAGGTTGGAGCCACAGACACCCACCGCCACCCCGGCTG 1644
1922 GGTCTGGCTCTCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1981
1645 GGTCTGGCTCTCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1704
1982 CCCCCCTATGAGGCGGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAG 2041
1705 CCCCCCTATGAGGCGGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAGGCTCTGAG 1764
2042 CACACCCATGGTTTTCAATTTGACCGCCCGCTCTCTGCTCATATGACACCCAGCTCT 2101
1765 CACACCCATGGTTTTCAATTTGACCGCCCGCTCTCTGCTCATATGACACCCAGCTCT 1824
2102 TTGAGGAGATCAGAGCCCATTTGCAAGAGAGCGGCTGCTGCAACCATCTTGTCTCCGA 2161
1825 TTGAGGAGATCAGAGCCCATTTGCAAGAGAGCGGCTGCTGCAACCATCTTGTCTCCGA 1884
2162 TTGCAAAATGACACCCCGATCTAGACATCTCTCAAGCCCTTTAACTCAGATGTCAA 2221
1885 TTGCAAAATGACACCCCGATCTAGACATCTCTCAAGCCCTTTAACTCAGATGTCAA 1944
2222 GCCACCGGCAACCCCGCTCAATACCTCCACCAAGAGATGATGAGACCTTCACTG 2281
1945 GCCACCGGCAACCCCGCTCAATACCTCCACCAAGAGATGATGAGACCTTCACTG 2004
2282 CTCGCCCAACCCAGCTGAGGTGGGACAGCGCAACGCTGTTCCGGTTGGAAACAGCAGA 2341
2005 CTCGCCCAACCCAGCTGAGGTGGGACAGCGCAACGCTGTTCCGGTTGGAAACAGCAGA 2064
2342 GGTCTCAGAACTGGCTCTGAATAGGACAGACCTAGCAGAGGAGATCAGGGTATCGGG 2401
2065 GGTCTCAGAACTGGCTCTGAATAGGACAGACCTAGCAGAGGAGATCAGGGTATCGGG 2124
2402 CGTTTGAAGTTTTCAGAGTCAATTCGGGAGATTAATCCAGTGGCTGGCCCGCAGCCACC 2461
2125 CGTTTGAAGTTTTCAGAGTCAATTCGGGAGATTAATCCAGTGGCTGGCCCGCAGCCACC 2184
2462 TGCATTCAGAGTCTGCAAGAGTCTTGTGGGAGGCAAAATTCCTAGGAAAG 2521
2185 TGCATTCAGAGTCTGCAAGAGTCTTGTGGGAGGCAAAATTCCTAGGAAAG 2244
2522 AAGACAGATTTTCTAATGTGTCCAAATGCGGATCAGTGTGATGACCTCTAGAGC 2581
2245 AAGACAGATTTTCTAATGTGTCCAAATGCGGATCAGTGTGATGACCTCTAGAGC 2304
2582 ACTGAGTCTCCCTGCTCTGCAAGTATTTAAGAAAGGCTGGCCAGGACGATGCTCAC 2641
2305 ACTGAGTCTCCCTGCTCTGCAAGTATTTAAGAAAGGCTGGCCAGGACGATGCTCAC 2364
2642 GCCTGTAATCCAGACTTTGGGAGGCGGAGGCGGATCACTGAGTGGAGGTTG 2701
2365 GCCTGTAATCCAGACTTTGGGAGGCGGAGGCGGATCACTGAGTGGAGGTTG 2424
2702 AAGACAGCTTGGCCCAATGTTGAACCTCTCTCTAATAAATACAAAATTAGCCAG 2761
2425 AAGACAGCTTGGCCCAATGTTGAACCTCTCTCTAATAAATACAAAATTAGCCAG 2484
2762 GGTGTGTGAGCTGCTGTAATCCAGCTTCTGGAGGCTGAGGATGAAATCACTT 2821
2485 GGTGTGTGAGCTGCTGTAATCCAGCTTCTGGAGGCTGAGGATGAAATCACTT 2544
2822 AAACCTGAGGAGGAGGTTTACAGTGAGCCAGATCGTGCCACTGCACTGAGGCTGGGC 2881

Db 2545 AAACCTGAGGAGGAGGTTTACAGTGAGCCAGATCGTGCCACTGATCCAGCCTGGGC 2604
Qy 2882 GACAGAGCAAGACTCTGTCTC 2902
Db 2605 GACAGAGCAAGACTCTGTCTC 2625
RESULT 13
ADRI6274
ID ADRI6274 standard; cDNA; 2166 BP.
XX AC ADRI6274;
XX DT 21-OCT-2004 (first entry)
XX DE Mouse jittery cDNA.
KW Mouse; jittery; ss; gene; Cayman ataxia; ATCAY; KIAA1872; chromosome 10;
XX ataxia; myoclonus; dystonia; epilepsy; nyctagmus.
OS Mus musculus.
FH Key Location/Qualifiers
CDS 378..1496
FT /*tag= a
FT /product= "Jittery"
FT /transl_except= (pos:1044..1046,aa:Glu)
XX US2004146900-A1.
XX 29-JUL-2004.
XX 03-NOV-2003; 2003US-00699941.
XX 01-NOV-2002; 2002US-0422971P.
PR 08-NOV-2002; 2002US-0424973P.
XX (UNMI) UNIV MICHIGAN.
XX Burmeister M;
XX WPI; 2004-552665/53.
DR P-PSDB; ADRI6275.
XX Detecting variant Cayman ataxia polypeptide or nucleic acid sequence in
subject, by detecting presence or absence of variant Cayman ataxia
polypeptide or nucleic acid in biological sample.
XX Disclosure; SEQ ID NO 1; 98pp; English.
CC The invention relates to detecting variant Cayman ataxia polypeptide or
nucleic acid sequence (ATCAY) in a subject, involving providing a
biological sample from a subject, where the biological sample comprises a
Cayman ataxia polypeptide or nucleic acid, and detecting the presence or
absence of a variant Cayman ataxia polypeptide or nucleic acid in the
biological sample. Also included is a kit comprising a reagent for
detecting the presence or absence of a variant Cayman ataxia nucleic acid
or polypeptide in a biological sample. The human ATCAY gene (previously
isolated as cDNA KIAA1872, located on chromosome 19p13.3) was identified
and mapped using the fact that it is the homologue of the mouse Jittery
gene, located on mouse chromosome 10. The method is useful for detecting
variant Cayman ataxia polypeptide or nucleic acid sequence in a subject.
CC The presence of the variant Cayman ataxia polypeptide or nucleic acid is
indicative of a disorder chosen from ataxia, myoclonus, dystonia,
epilepsy, and nyctagmus in the subject. The biological sample is chosen
from blood sample, a tissue sample, urine sample, saliva sample, and an
amniotic fluid sample. The subject is chosen from embryo, foetus, newborn
animal, young animal, and an adult animal. The animal is a human. The
human is an adult female of child-bearing age. The present sequence is
the cDNA encoding the mouse Jittery protein.
XX Sequence 2166 BP; 523 A; 619 C; 576 G; 448 T; 0 U; 0 Other;

[illegible]

PT cDNAs.

Claim 1; SEQ ID NO 5423; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH18632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 790 BP; 166 A; 245 C; 251 G; 125 T; 0 U; 3 Other;

Query Match 25.5%; Score 745.2; DB 4; Length 790;
Best Local Similarity 99.2%; Pred. NO. 1.5e-171;
Matches 758: Conservative 0; Mismatches 5; Indels 1

Qy	8	CTCTGCGACCCCTGAGCTTGGGAAGACAGCTACCTCGAGAGCAGGGCGCGCAGCGGGCG	67
Db	1	CTCTGCCACGCCCCCGGGCTTGGGAAGACAGCTACCTCGAGAGCAGGGCGCGCAGCGGGCG	60
Qy	68	GGCGATGAGAGGGGGCGCAGCGCAGCCCGCGCTGGGGAGCCCCACCGCTAACCCCTGCAC	127
Db	61	GGCGATGAGAGGGGGCGCAGCGCAGCCCGCGCTGGGGAGCCCCACCGCTAACCCCTGCAC	120
Qy	128	CCCAACCCACCCCTGCAAAAAGAGTGGCGGGCGCTGGCCACGTCGCGCTTGGGTGACCTT	187
Db	121	CCCAACCCACCCCTGCAAAAAGAGTGGCGGGCGCTGGCCACGTCGCGCTTGGGTGACCTT	180
Qy	188	CCTCGGATGCAGAAATCGCGCCCTCGGAGCATCTCTTCTCTCTAGGCTCTBAAGGCCCG	247
Db	181	CCTCGGATGCAGAAATCGCGCCCTCGGAGCATCTCTTCTCTCTAGGCTCTBAAGGCCCG	240
Qy	248	GGAGCGTGAGCGATGCCCGCAGCTGCACCGGGCAGGGCTCGCTTCAAGCCAGTGCCTTCC	307
Db	241	GGAGCGTGAGCGATGCCCGCAGCTGCACCGGGCAGGGCTCGCTTCAAGCCAGTGCCTTCC	300
Qy	308	GGAGAGGCTGTCTCAGCTGCGAGAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTTCCCGAC	367
Db	301	GGAGAGGCTGTCTCGGCTGCGAGAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTTCCCGAC	360
Qy	368	CTCCCATCGGGACCAACCGAAGCCACGCTCGGATGGAACCGTGGACGTGAAGAGGAAT	427
Db	361	CTCCCATCGGGACCAACCGAAGCCACGCTCGGATGGAACCGTGGACGTGAAGAGGAAT	420
Qy	428	GGCAGGACGAAGATCTTCCAGGCGCATCCCGAAGAGACGGGGTGGAACTGCTTGGCA	487
Db	421	GGCAGGACGAAGATCTTCCAGGCGCATCCCGAAGAGACGGGGTGGAACTGCTTGGCA	480
Qy	488	GCCCGGTGGAAAGACATCTCTCTCTCCCAACCGCTAAATTTCAACGGAGCGCATCGTA	547
Db	481	GCCCGGTGGAAAGACATCTCTCTCTCCCAACCGCTAAATTTCAACGGAGCGCATCGTA	540
Qy	548	AGAGGAAGACGCTGGTGGCCCCACAGATCAACATTTCTCTGATACAGATGAGGGGTGCC	607
Db	541	AGAGGAAGACGCTGGTGGCCCCACAGATCAACATTTCTCTGATACAGATGAGGGGTGCC	600

SQ Sequence 703 BP; 152 A; 218 C; 213 G; 120 T; 0 U; 0 Other;
Query Match 23.1%; Score 675; DB 5; Length 703;
Best Local Similarity 98.4%; Pred. No. 2e-154;
Matches 692; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
QY 66 GCGGCGATGAGAGGGGCGAGCGGAGCCCGCGCTGGGAGCCACCGCTAACCCCTGC 125
DB 1 GCAGAGATGAGAGGGGCGAGGCTCAGCCCGCGCTGGGAGCCACCGCTAACCCCTGC 60
QY 126 ACCCCACCCACCCCTGCACAAAAGAGCTGGCGGGCGCTGCCACGTCGCCCTGGGGTGACC 185
DB 61 ACCCCACCCACCCCTGCACAAAAGAGCTGGCGGGCGCTGGCCACGTCGCCCTGGGGTGACC 120
QY 186 TTCCTCGGATGCAAGATCCGCCCTCGAGCATCTCTTCCTCTAGGCTCTGAAGGCC 245
DB 121 TTCCTCGGATGCAAGATCCGCCCTCGAGCATCTCTTCCTCTAGGCTCTGAAGGCC 180
QY 246 GGGGAGGTGAGCGATGCCAGCTGCACCGGGCAGGGCTCGCCTTTGTTGGCA-GTAA 304
DB 181 GGGGAGGTGAGCGATGCCAGCTGCACCGGGCAGGGCTCGCCTTTGTTGGCATGTAA 240
QY 305 GAGGAGAGGCTCTCTCAGCTGCAGAGGGGTCACTCCCTCAAGCCAGTGCCTCTCC 364
DB 241 GAGGAGAGGCTCTCTCAGCTGCAGAGGGGTCACTCCCTCAAGCCAGTGCCTCTCC 300
QY 365 CAGCTCCCATGGGGACCAACCGAAGCCAGCTCCGGATGGAAACGTGGACGTGAAGGAG 424
DB 301 CAGCTCCCATGGGGACCAACCGAAGCCAGCTCCGGATGGAAACGTGGACGTGAAGGAG 360
QY 425 AATGGCAGGACGAAGATCTTCCAGGCGCACTCCAGAGAGAGCGGGGTGGAACCTGCTTG 484
DB 361 AATGGCAGGACTAAGATCTTCCAGGCGCACTCCAGAGAGAGCGGGGTGGAACCTGCTTG 420
QY 485 GCAGCCCGTGGAGAGACACATCTCTCCCAACAGCTTAATTTCAACGGAGCGCATC 544
DB 421 GCAGCCCGTGGAGAGACACATCTCTCCCAACAGCTTAATTTCAACGGAGCGCATC 480
QY 545 GTAAGAGGAGACGCTGGTGGCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAGGGGT 604
DB 481 GCAAGAGGAGACGCTGGTGGCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAGGGGT 540
QY 605 CCCTGCTGTCGATGACTTCTGGATACCCCTGATGACCTGGATATTAACGTGGATGACA 664
DB 541 CCCTGCTGTCGATGACTTCTGGATACCCCTGATGACCTGGATATTAACGTGGATGACA 600
QY 665 TCGAGACCCCGATGAGACCGACTCGCTGGAGTTCCTGGGGATGGCAACGAACCTGGACT 724
DB 601 TCGAGACCCCGATGAGACCGACTCGCTGGAGTTCCTGGGGATGGCAACGAACCTGGACT 660
QY 725 GGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGG 767
DB 661 GGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGG 703

Search completed: March 5, 2006, 03:30:55
Job time : 1683 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2887.6	99.0	3076	3	US-09-174-937-1	Sequence 1, Appli
2	2598.2	89.0	2759	3	US-10-104-047-997	Sequence 997, App
3	1325.8	45.4	2625	3	US-10-104-047-986	Sequence 986, App
4	1113	38.1	1113	3	US-09-174-937-3	Sequence 3, Appli
5	309.8	10.6	322	3	US-09-621-978-14512	Sequence 14512, A
6	302.2	10.4	966	3	US-09-174-937-6	Sequence 6, Appli
7	302.2	10.4	4235	3	US-09-174-937-4	Sequence 4, Appli
8	241.8	8.3	103987	3	US-09-949-016-12513	Sequence 12513, A
9	241.8	8.3	103988	3	US-09-949-016-17050	Sequence 17050, A
10	241.2	8.3	33712	3	US-09-949-016-15793	Sequence 15793, A
11	241.2	8.3	193169	3	US-09-949-016-15091	Sequence 15091, A
12	240.4	8.2	87734	3	US-09-949-016-17521	Sequence 17521, A
13	239.8	8.2	49378	3	US-09-949-016-13408	Sequence 13408, A
14	239.6	8.2	11929	3	US-09-949-016-17290	Sequence 17290, A
15	239.4	8.2	24221	3	US-09-949-016-14964	Sequence 14964, A
16	239.2	8.2	21590	3	US-09-949-016-13372	Sequence 13372, A
17	239.2	8.2	21590	3	US-09-949-016-13373	Sequence 13373, A
18	239	8.2	601	3	US-09-949-016-183967	Sequence 183967, A
19	239	8.2	601	3	US-09-949-016-184072	Sequence 184072, A
20	239	8.2	601	3	US-09-949-016-184177	Sequence 184177, A
21	239	8.2	601	3	US-09-949-016-184282	Sequence 184282, A
22	239	8.2	43543	3	US-09-778-961-3	Sequence 3, Appli
23	239	8.2	246230	3	US-09-949-016-17019	Sequence 17019, A
24	239	8.2	246230	3	US-09-949-016-17020	Sequence 17020, A

Db	321	CTGCAGAGGGGTCACTCCCTGCTTCAGGCCAGTGCCTCTTCCCAAGCTCCCATGGGACAC	380
Qy	384	CGAAGCCACGCTCCGGATGGAAAAACGTGGACGCTGAAGAGAGGAATGGCAGACGAGAATCT	443
Db	381	CGAAGCCACGCTCCGGATGGAAAAACGTGGACGCTGAAGAGAGGAATGGCAGACGAGAATCT	440
Qy	444	TCCCAGGCCACTCCCAGAAGAGACGGGGGTGGAACTGCTTGGCAGCCCGGTGGAAAGACAC	503
Db	441	TCCCAGGCCACTCCCAGAAGAGACGGGGGTGGAACTGCTTGGCAGCCCGGTGGAAAGACAC	500
Qy	504	ATCCTCTCTCCCAACACAGCTAAATTTCAACGGAGCGCATCGTAAGAGAGAGACGCTGGT	563
Db	501	ATCCTCTCTCCCAACACAGCTAAATTTCAACGGAGCGCATCGTAAGAGAGAGACGCTGGT	560
Qy	564	GGCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAGGGGTCTCTGCTGCTCGATGACTT	623
Db	561	GGCCCCAGAGATCAACATTTCTCTGGATCAGAGTGAGGGGTCTCTGCTGCTCGATGACTT	620
Qy	624	CTTGGATACCCCTGATGACCTGGATATTAACCTGGATGACATCGAGACCCCGCATGAGAC	683
Db	621	CTTGGATACCCCTGATGACCTGGATATTAACCTGGATGACATCGAGACCCCGCATGAGAC	680
Qy	684	CGACTCGCTGAGTTCCTGGGGAATGGAAACGAACTCGAGTGGGAAGACGACACCCCGT	743
Db	681	CGACTCGCTGAGTTCCTGGGGAATGGAAACGAACTCGAGTGGGAAGACGACACCCCGT	740
Qy	744	GGCCACCGCCAAAGAAATGCCCGGGGACGGCGGATCTATTTGGGACGGCAGCGGA	803
Db	741	GGCCACCGCCAAAGAAATGCCCGGGGACGGCGGATCTATTTGGGACGGCAGCGGA	800
Qy	804	GGACGGCAGCGCCGCAACCGGGCCCTGTGGCGGACAGTGATCATCGGAGAGCAAGACA	863
Db	801	GGACGGCAGCGCCGCAACCGGGCCCTGTGGCGGACAGTGATCATCGGAGAGCAAGACA	860
Qy	864	CCGTATAGACTGCACATGATCCGGCCTTATCATGAAAGTGGTCAACCACGGAGGGTACTA	923
Db	861	CCGTATAGACTGCACATGATCCGGCCTTATCATGAAAGTGGTCAACCACGGAGGGTACTA	920
Qy	924	CGCGAAGGCTCAACGCCATCATCGTCTTGGCAGCCTGCTTCTTCCAGCAGCAGCT	983
Db	921	CGCGAAGGCTCAACGCCATCATCGTCTTGGCAGCCTGCTTCTTCCAGCAGCAGCT	980
Qy	984	CCCCGACTACCACTACATCATGGAGAACCTCTTCTGTAGTCAATCAGCAGCTTAGAGCT	1043
Db	981	CCCCGACTACCACTACATCATGGAGAACCTCTTCTGTAGTCAATCAGCAGCTTAGAGCT	1040
Qy	1044	CCTGGTGGCTGAGGACTACATGATCGTGTACCTGGAACGGTGCCACGCCCGGGAGGAT	1103
Db	1041	CCTGGTGGCTGAGGACTACATGATCGTGTACCTGGAACGGTGCCACGCCCGGGAGGAT	1100
Qy	1104	GCCTGGAATCGGCTGGAGTGAAGAGTGCTACACAGATGATCGACGGAGGTTCGGAAGAA	1163
Db	1101	GCCTGGAATCGGCTGGAGTGAAGAGTGCTACACAGATGATCGACGGAGGTTCGGAAGAA	1160
Qy	1164	CCTGAACTCTCTGATCATCGTCCACCCCTCGTGGTTCAATTCGAGCTGTGCTGGCCATCTC	1223
Db	1161	CCTGAACTCTCTGATCATCGTCCACCCCTCGTGGTTCAATTCGAGCTGTGCTGGCCATCTC	1220
Qy	1224	TCGCCCCCTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTACGTGCAACAGCTTGGAA	1283
Db	1221	TCGCCCCCTTTCATCAGCGTCAAGTTTCATCAACAGATCCAGTACGTGCAACAGCTTGGAA	1280
Qy	1284	CCTGGAGCAACTCATCCCTATGGAAACAGTCCAGATCCAGATCTGCGTCTCGAATACGA	1343
Db	1281	CCTGGAGCAACTCATCCCTATGGAAACAGTCCAGATCCAGATCTGCGTCTCGAATACGA	1340
Qy	1344	AGAGAAAGACTGAAGCCAGGAGGGAGAGCGGAGGCCCGAGGCCCGAGTTTGTGCTGCC	1403
Db	1341	AGAGAAAGACTGAAGCCAGGAGGGAGAGCGGAGGCCCGAGGCCCGAGTTTGTGCTGCC	1400
Qy	1404	CAGGTCTGAAGAGAACGAGGTGGCACCATGTGAAAAACAGGTCTCTCTGGTCTCAGA	1463

QY 1388 CGAGTTTGTGCTGCCAGTCTGAAGAGAGCCAGAGGTGGCAACAGTGGAAAAAGAGT 1447
| | | | |
Db 1245 CGAGTTTGTGCTGCCAGTCTGAAGAGAGCCAGAGGTGGCAACAGTGGAAAAAGAGT 1304
| | | | |
QY 1448 CTGCTCTGTGCTCAGAGATCAGAAACAGCATGTCTGAGGCGAGGTGAGATACAA 1507
| | | | |
Db 1305 CTGCTCTGTGCTCAGAGATCAGAAACAGCATGTCTGAGGCGAGGTGAGATACAA 1364
| | | | |
QY 1508 AGGACATGGAAGAGATTCAGATGCGCAGAAAACTCTGTGAGAGCCCACTGGCCCCAG 1567
| | | | |
Db 1365 AGGACATGGAAGAGATTCAGATGCGCAGAAAACTCTGTGAGAGCCCACTGGCCCCAG 1424
| | | | |
QY 1568 ATCTCATCTGCTCATCTGAGTCCCAATCTTCCAAAGGTGCGAGCCCTCGTTTCATC 1627
| | | | |
Db 1425 ATCTCATCTGCTCATCTGAGTCCCAATCTTCCAAAGGTGCGAGCCCTCGTTTCATC 1484
| | | | |
QY 1628 TCTGAAACCCAGCATCTTTTACGCTGCTGAAAAATTTGTAATTTTTTTTTTTTAAAGAT 1687
| | | | |
Db 1485 TCTGAAACCCAGCATCTTTTACGCTGCTGAAAAATTTGTAATTTTTTTTTTTTAAAGAT 1544
| | | | |
QY 1688 GCAGTATTTGTGGTTTCCAGAAAGGCGCCAGCTCTGAGCCCTCACCCCTTCCACACTCA 1747
| | | | |
Db 1545 GCAGTATTTGTGGTTTCCAGAAAGGCGCCAGCTCTGAGCCCTCACCCCTTCCACACTCA 1604
| | | | |
QY 1748 CGAATCTCAGCGAGGAAGCAAGAGCGAGGGGTGGCCCGGTGGCGTGGTGGCC 1807
| | | | |
Db 1605 CGAATCTCAGCGAGGAAGCAAGAGCGAGGGGTGGCCCGGTGGCGTGGTGGCC 1664
| | | | |
QY 1808 TCCGCTCTGCTCGAGCCCTCTGCTGAGCTGAGTGGATACAGATTTCAAGCCCTTCTCT 1867
| | | | |
Db 1665 TCCGCTCTGCTCGAGCCCTCTGCTGAGCTGAGTGGATACAGATTTCAAGCCCTTCTCT 1724
| | | | |
QY 1868 TGCTTGTACCCCTCCAGTTTGGAGCCACAGACACCCAGCCGACCCGGCTGGGTCTG 1927
| | | | |
Db 1725 TGCTTGTACCCCTCCAGTTTGGAGCCACAGACACCCAGCCGACCCGGCTGGGTCTG 1784
| | | | |
QY 1928 GTGCTTTTCTGTGCTTCCCTCCAGAAATGCGGCTCAGACCTAGAACTCAACCCCC 1987
| | | | |
Db 1785 GTGCTTTTCTGTGCTTCCCTCCAGAAATGCGGCTCAGACCTAGAACTCAACCCCC 1844
| | | | |
QY 1988 TATGAGGGCCAGCTCTGGGTAGCTCTGACCTCGACCTTATGTCCAAATTCACCC 2047
| | | | |
Db 1845 TATGAGGGCCAGCTCTGGGTAGCTCTGACCTCGACCTTATGTCCAAATTCACCC 1904
| | | | |
QY 2048 CATGGTTTTTCAATTTGACCCGCCCTTCTGCTCATATGACACCCAGCTCTTTGAGA 2107
| | | | |
Db 1905 CATGGTTTTTCAATTTGACCCGCCCTTCTGCTCATATGACACCCAGCTCTTTGAGA 1964
| | | | |
QY 2108 GGATCAGAGCCCATTTGCAAGAGAGCGCTGCCAAACCATCTTGTCTCGAATTCGAA 2167
| | | | |
Db 1965 GGATCAGAGCCCATTTGCAAGAGAGCGCTGCCAAACCATCTTGTCTCGAATTCGAA 2024
| | | | |
QY 2168 AATGACACCCAGTATCTAGACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCAC 2227
| | | | |
Db 2025 AATGACACCCAGTATCTAGACATTTCTCAAGCCCTTTAACTCAGATGTCAAGCCAC 2084
| | | | |
QY 2228 GGGCAAAACCCGTCATACCTCCCAACAGGAATGAGATATGTGGACCTCACTGTCCCC 2287
| | | | |
Db 2085 GGGCAAAACCCGTCATACCTCCCAACAGGAATGAGATATGTGGACCTCACTGTCCCC 2144
| | | | |
QY 2288 CAACCCAGGCTCAGGCTGGGACACGCAACGCTGTTCCGGGTTGGAACAGCAGAGCTCA 2347
| | | | |
Db 2145 CAACCCAGGCTCAGGCTGGGACATGCAACGCTGTTCCGGGTTGGAACAGCAGAGCTCA 2304
| | | | |
QY 2348 GAACTGGCTCTCAAAATGCGACCTTACAGAGGAGATACAGGATTCGGGCTTTG 2407
| | | | |
Db 2205 GAACTGGCTCTCAAAATGCGACCTTACAGAGGAGATACAGGATTCGGGCTTTG 2264
| | | | |
QY 2408 AGTGTTCAGAGTCAATTCGGGAAGATTAATCAGGTGCGTGGCCGCGCAGCTGCATT 2467
| | | | |
Db 2265 AGTGTTCAGAGTCAATTCGGGAAGATTAATCAGGTGCGTGGCCGCGCAGCTGCATT 2324
| | | | |
QY 2468 CAAAGCTTGGACCGAGGGTTCTTTGTTCCGGGAGGCAAAATTTCCCTTAGGAAAAAGAGACA 2527
| | | | |

RESULT 3

US-10-104-047-986

; Sequence 986, Application US/10104047

; Patent No. 6943241

; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE

; TITLE OF INVENTION: No. 6943241el full length cdna

; FILE REFERENCE: HL-A0105

; CURRENT APPLICATION NUMBER: US/10/104,047

; PRIOR FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER:

; NUMBER OF SEQ ID NOS: 4096

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 986

; LENGTH: 2625

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-104-047-986

Query Match 45.4%; Score 1325.8; DB 3; Length 2625;

Best Local Similarity 99.8%; Pred. No. 8.7e-314;

Matches 1338; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1562 CCCAGATCTCATCTGCTCATCTGCTCATCTGCTCCCAATCTTCCAAAGGTGCGAGCCCTCG 1621
| | | | |
Db 1286 CTCAGATCTCATCTGCTCATCTGCTCATCTGCTCCCAATCTTCCAAAGGTGCGAGCCCTCG 1345
| | | | |
QY 1622 TTCACTCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAAGAAATTTGTAATTTTTTTTTTT 1681
| | | | |
Db 1346 TTCACTCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAAGAAATTTGTAATTTTTTTTTTT 1404
| | | | |
QY 1682 AACGATGCGATATTTGTGGTTTCCAGAAAGGCGCCAGCTCTGAGCCCTCACTCCCTTCCA 1741
| | | | |
Db 1405 AACGATGCGATATTTGTGGTTTCCAGAAAGGCGCCAGCTCTGAGCCCTCACTCCCTTCCA 1464
| | | | |
QY 1742 CACTCAGCAACTCTCAGCCGAGGAGCAAGAGCGAGGGGTGGCCGCGTGGCGTGG 1801
| | | | |
Db 1465 CACTCAGCAACTCTCAGCCGAGGAGCAAGAGCGAGGGGTGGCCGCGTGGCGTGG 1524
| | | | |
QY 1802 GTGGCTCGCTCTCTGCTCGCAGCCCTCTGTGTGTGAGCTGGATACAGATTTCAAGACCC 1861
| | | | |

QY 1033 AGCTTAGAGTCTCTGGTGGCTGAGGACTACATGATCGTGTACTTGAAACGGTGCACGCC 1092
Db 661 AGCTTAGAGTCTCTGGTGGCTGAGGACTACATGATCGTGTACTTGAAACGGTGCACGCC 720
QY 1093 CGCGGAGGATGCTGGAATCGGCTGGCTGAGAGTGTACAGATGATCGACCGGAGG 1152
Db 721 CGCGGAGGATGCTGGAATCGGCTGGCTGAGAGTGTACAGATGATCGACCGGAGG 780
QY 1153 TTGCGGAAACCTGGAAGTCTTGATCATCGTCCACCCCTCGTGGTTCAATTCGACTGTG 1212
Db 781 TTGCGGAAACCTGGAAGTCTTGATCATCGTCCACCCCTCGTGGTTCAATTCGACTGTG 840
QY 1213 CTGGCCATCTCTGGCCCTTTCATCAGCGTCAAGTTTCATCAACAAGATCCAGTACGTGCAC 1272
Db 841 CTGGCCATCTCTGGCCCTTTCATCAGCGTCAAGTTTCATCAACAAGATCCAGTACGTGCAC 900
QY 1273 AGCTTGGAGACCTGGAGCACTCATCCCTATGGAACAGTCCAGATCCAGACTCGTC 1332
Db 901 AGCTTGGAGACCTGGAGCACTCATCCCTATGGAACAGTCCAGATCCAGACTCGTC 960
QY 1333 CTGCAATACGAAGAGAAAGACTGAAGGCGAGGAGGAGCGCGAGGCCCGCGGAG 1392
Db 961 CTGCAATACGAAGAGAAAGACTGAAGGCGAGGAGGAGCGCGAGGCCCGCGGAG 1020
QY 1393 TTGTGCTGCCAGGCTCTGAAGAGAGCCAGAGGTGCGCACCAAGTGGAAACAGGTTCTGT 1452
Db 1021 TTGTGCTGCCAGGCTCTGAAGAGAGCCAGAGGTGCGCACCAAGTGGAAACAGGTTCTGT 1080
QY 1453 CTGCTCTCAGAGATCAGGAAACAGCATGTCC 1485
Db 1081 CTGCTCTCAGAGATCAGGAAACAGCATGTCC 1113

RESULT 5

US-09-621-976-14512
; Sequence 14512, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 14512
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-14512

Query Match 10.6%; Score 309.8; DB 3; Length 322;
Best Local Similarity 98.8%; Pred. No. 4.8e-66;
Matches 319; Conservative 3; Mismatches 0; Indels 1; Gaps 1;
QY 2575 TAGAAGCACTGAGCTCCCTGCTCTCTGGAAGTATTTAAGAAAGGCTGGCGGACGACGAT 2634
Db 1 TAGAAGCACTGAGCTCCCTGCTCTCTGGAAGTATTTAAGAAAGGCTGGCGGACGACGAT 60
QY 2635 GGCTACGCGCTGTAATCCAGACTTTTGGGAGGCGGAGGAGGCGGATCACTGAGGTGAG 2694
Db 61 GGCTACGCGCTGTAATCCAGASTTTTGGGAGGCGGAGGAGGCGGATCACTGAGGTGAG 120
QY 2695 GAGTTTGAGAACAGCTCGGCAACATGGTGGAACCTCTCTCTAATAAATAACAAAAT 2754
Db 121 GAGTTTGAGAACAGCTCGGCAACATGGTGGAACCTCTCTCTAATAAATAACAAAAT 180
QY 2755 TAGCCAGGCGTGGTGGCAGTGTCTTAATCCAGCTACTTGGAGGCTGAGGCATGAGA 2814
Db 181 TAGCCAGGCGTGGTGGCAGTGTCTTAATCCAGCTACTT-GGAGGCTGAGGCATGAGA 239

QY 2815 ATCACTTAAACCTGAGAGCGAGAGGTTACAGTGAGCCAGAGATCGTGCCACTGCAATTCAG 2874
Db 240 ATCACTTAAACCTGAGAGCGAGAGGTTACAGTGAGCCAGAGATCGTGCCACTGCAATTCAG 299
QY 2875 CCTGGCGACAGAGCAAGACTCT 2897
Db 300 CCTGGCGACAGAGCAAGACTCT 322

RESULT 6

US-09-174-937-6
; Sequence 6, Application US/09174937
; Patent No. 6756212
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: Isolated Proteins and Nucleic Acid Molecules Having
; TITLE OF INVENTION: Homology to NIP2 Proteins and Uses Thereof
; FILE REFERENCE: mm1061
; CURRENT APPLICATION NUMBER: US/09/174,937
; CURRENT FILING DATE: 1998-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(966)
US-09-174-937-6

Query Match 10.4%; Score 302.2; DB 3; Length 966;
Best Local Similarity 62.5%; Pred. No. 5.2e-64;
Matches 547; Conservative 0; Mismatches 28; Indels 45; Gaps 3;
QY 469 GGGGTGGAACTGCTTTGGCAGCCCGGTGGAAGACACATCTCTCTCTCCCAACACGCTAAAT 528
Db 97 GAGGGCGTGTGAGTCCCGAGTGTGCAGACATGAGGCTTGAACCTCTTAATTTCTCTGGAT 156
QY 529 TTCACGCGAGCGCATCGTAAGAGGAGAGCGTGTGGTGGCCCGCCAGAGATCAACATTTCTCTG 588
Db 157 CTTAATGACATCTCATCTCTCGGAGAAATCAAGTGTCAAGCCCAAAATATCAATCTTTCTCTG 216
QY 589 GATCAGATGAGGGTCCCTGCTGCCATGACTTCTTGGATACCCCTGATGACCTGGAT 648
Db 217 GACCAAGTGAAGATCTATTCTCTGATGATTAATTTGGACAGTCCAGATGAATTCAC 276
QY 649 ATTAACTGGATGACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTCTGGGAAT 708
Db 277 ATCAATGTGGATGAATTTGATACCCCGATGAAGCAGATTTCTTTTGGTACACTGG--- 332
QY 709 GGCAAGCAACTGGAGTGGGAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGGG 768
Db 333 -----CCATGATCCCAAGCAACAAAGATTTCTGGC 363
QY 769 GACAGCGGGATCTATTTGGGAGCGGACGAGCGGAGGAGCGGAGCGCCCAACCGGGCGC 828
Db 364 CAAGAGTCAGAGTCTATTTCAGATATATACGGCCGAGAGAGGAAAGGAGAGCAAC---CG 420
QY 829 CTGTGGCGGACAGTGTATCATCGGGAGCAAGAGCAACCGTATAGACTCTGCATGATCCGG 888
Db 421 CTTTGGATGACAGTGGTCAITGGAGAACAGAGCAGCGCATTTGACATGAAGTTCATCGAG 480
QY 889 CCTTACATGAAGTGGTCAACCAAG-----AGGGTACTACGGCGAGGCGCTCAAC 939
Db 481 CCCTACAGGAGATCATTTTCTCACGGAGGAGATTGAGGATCTATGCGGAGCGGTCTAAAT 540
QY 940 GCATCATCTGCTTTCGCGAGCTGCTTCTTCCAGACAGCAGCTCCCGAGCTACCACTAC 999
Db 541 GCCATCATTTGTTGGCGCTGTTTCTGCCAGACAGCAGTCCGGCGGATTAACACTAT 600
QY 1000 ATCATGGAGAACCTTCTCTGTAGCTCATCAGCAGCTTTAGAGCTCTCTGGTGGCTGAGGAC 1059

Db 601 GTCATGGAAATCTTTTCTATATATTAATAGTACTTTAGAGTTGATGTTAGCTGAAGAC 660
QY 1060 TACATGATCGTGTACGAAACGTCACGCGCCCGCGGAGGATGCTTGGAAATCGGCTGG 1119
Db 661 TATATGATGTTGTTGATGTCGCAACCCCAAGAGGAGGATGCCAGGGCTAGGCTGG 720
QY 1120 CTGAAGAAGTCTTACAGATGATCGACGGAGGTTCCGGAAGAAACCTGAAGTCTTGTATC 1179
Db 721 ATGAAGAAATGCTACAGATGATTTGACAGACGGTTGAGGAAGAAATTTGAATCATTTTCATC 780
QY 1180 ATCGTCCACCCCTCGTGGTTTCATTGCGACTGTGTCGCCATCTCTCGCCCTTTTCATCAGC 1239
Db 781 ATGTTTCATCCATCTTGGTTTCATCAGAACATCTTGTGTCAGACGACCTTTTATAAGT 840
QY 1240 GTCAAGTTCATCAACAGATCCAGTACGTGTCAGAGCTTGGAGAGACCTGGAGCAACTCATC 1299
Db 841 TCAGAAATTCAGCAGTAAATTAATATGTCATATGTCATATCAGAACTCAGTGGGCTGATC 900
QY 1300 CCTATGGACAGCTCCAGATCCAGACTCGGTCT 1334
Db 901 CCAATGGATTGCATCCACATTCAGAGAGCATCAT 935

RESULT 7

US-09-174-937-4

; Sequence 4, Application US/09174937

; Patent No. 6756212

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran

; TITLE OF INVENTION: Isolated Proteins and Nucleic Acid Molecules Having

; TITLE OF INVENTION: Homology to NIP2 Proteins and Uses Thereof

; FILE REFERENCE: mn1061

; CURRENT APPLICATION NUMBER: US/09/174,937

; CURRENT FILING DATE: 1998-10-19

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 4

; LENGTH: 4235

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (23)..(988)

US-09-174-937-4

Query Match 10.4%; Score 302.2; DB 3; Length 4235;
Best Local Similarity 62.5%; Pred. No. 9.3e-64;
Matches 547; Conservative 0; Mismatches 283; Indels 45; Gaps 3;

QY 469 GGGGTGGAACTGCTTGGCAGCCCGGTGGGAAGACACATCTCTCTCCCAACACGCTAAAT 528
Db 119 GAGGCGTCTGAGTCCAGTCTGTCAGACATGAGGCTGACCTTCTTAATCTCTGAT 178
QY 529 TTCAACGAGCGGATCGTAAGAGGAAGACGCTGGTGGCCCAAGATCAACATTTCTCTG 588
Db 179 CTTAATGACACTCATCTCGGAGATCAAGCTCACAGCCCAATATCAATCTTCTCTG 238
QY 589 GATCAGAGTGGGGTCCCTGCTGTCGATGACTTCTTGATACCCCTGATCACTGAT 648
Db 239 GACCAAGTGAAGGATCTATTCTCTGATGATACTTGGACAGCTCCAGATGAATGAC 298
QY 649 ATTAACGTGGATGACATCGAGACCCCGGATGAGACCGACTCGCTGGAGTTCCTGGGAAT 708
Db 299 ATCAATGTGATGATGATGATATACCCCGGATGAGCAGATCTTTTGGTACTGCTG 354
QY 709 GGCACGAAGTGGAGTGGGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCGGG 768
Db 355 -----CCATGATCCCAAGCAACAAAGATTTGCG 385
QY 769 GACAGCGCGATCTATTGGGACGCGACGAGGAGCGGACGCGCCGCCAACCGGCGC 828
Db 386 CAAGAGTCAGATCTATTCCGAATATATACGCGCGGAGGAGGAAACGCGGAGGACAC 442

QY 829 CTGTGCGGACAGTGTATCATCGGGAGCAAGACCGGTATAGACTGTGCAATGATCGCG 888
Db 443 CTTTGGATGACAGTGGTCAATTGGAGAACAGAGCAGCGCATTTGACATGAAGTCTATCGAG 502
QY 889 CTTTACATGAAGTGGTCAACCCACGG-----AGGTACTACGGCGAAGGCTCTCAAC 939
Db 503 CCCTACAGGAGAGTCAATTTCTCACGGAGGAGATTTCAGGATACTATGGGACGGTCTAAAT 562
QY 940 GCATCATCTGCTTCGCGAGCGTCTTCTTCAGACAGCAGCGCTCCCGACTTACCACATAC 999
Db 563 GCATCATCTGCTTCGCGCGCTGTTTCTGCGACAGCAGCTCGGGCGGATTTACCACATAT 622
QY 1000 ATCATGGAAGACCTCTTCTCTGATCATCATGAGCTTTAGAGCTCTCTGTGGCTGAGGAC 1059
Db 623 GTCATGGAAATCTTTTCTATATGTAATAGTACTTTTAGAGTTGATGCTAGTCTGAAGAC 682
QY 1060 TACATGATCGTGTACTGAAACGCTGCGCGCCCGCGGAGGATGCTTGGAAATCGGCTGG 1119
Db 683 TATATGATGTTGTTGCTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAATGTTGTAAT 742
QY 1120 CTGAAGAAGTGTCTACAGATGATCGACCGGAGTTCGGAAGAAACCTGAAGTCTTGTATC 1179
Db 743 ATGAAGAAATGCTACAGATGATTCAGACAGCGTTGAGGAAGAAATTTGAATCATTTTCATC 802
QY 1180 ATCGTCCACCCCTCGTGGTTTCATTGCGACTGTGTCGGCCATCTCTCGCCCTTTTCATCAGC 1239
Db 803 ATTGTTTCATCCATCTTGGTTTCATCAGAACATCTTGTGTCAGACGACCTTTTATAAGT 862
QY 1240 GTCAAGTTCATCAACAGATCCAGTACGTGTCAGAGCTTGGGAAGACCTGGAGCAACTCATC 1299
Db 863 TCAGAAATTCAGCAGTAAATTAATATGTCATATGTCATATGTCATATGTCATATGTCATATGTCATAT 922
QY 1300 CCTATGGAAACAGCTCCAGATCCAGACTCGGTCT 1334
Db 923 CCAATGGATTGCATCCACATTCAGAGAGCATCAT 957

RESULT 8

US-09-949-016-12513/c

; Sequence 12513, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12513

; LENGTH: 103987

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-12513

Query Match 8.3%; Score 241.8; DB 3; Length 103987;
Best Local Similarity 89.2%; Pred. No. 1.8e-48;
Matches 272; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 2614 AAAGTCTGGGCGGAGGAGTGGCTTCAAGCTGTAATCCAG-ACCTTTGGGAGGCGGAGG 2672
Db 47908 AAAGTCTGGGCGGAGGAGTGGCTTCAAGCTGTAATCCAGCACTTTTGGGAGGCTGAGG 47849
QY 2673 CAGCGGATCACTTGGAGTGGAGGATTTGAGAACAGCCCTGGGCAACATGGTGAACCTCA 2732
Db 47848 CAGCGGATCACTTGGAGTGGAGGATTTGAGAACAGCCCTGGGCAACATGGTGAACCTCA 47789

|||||
Db 4403 AAAAGAAA 4395

RESULT 14

US-09-949-016-17290
; Sequence 17290, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17290
; LENGTH: 11929
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17290

Query Match 8.2%; Score 239.6; DB 3; Length 11929;
Best Local Similarity 87.3%; Pred. No. 2.7e-48;
Matches 274; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 2605 TATTTAAGAAAGCGCTGGGCGGCGAGATGGCTCAGCGCTGTAATCCCGAG-ACCTTTGGG 2663
Db 6427 TAAAAAATAAAAAATTTGGCGGGCGAGTGGCTCACCTGTAAATCCCGAGCACTTTGGG 6486

QY 2664 AGCCGAGGCGAGCGGATCACTTGAGGTGAGGAGTTTGAGAAAGAGCTGGCCCAACATGGT 2723
Db 6487 AGCCGAGGCGAGCGAGATCACTTGAAGTCAGGAGTTTCGAGACCAGCGCTGGCCCAACATGGT 6546

QY 2724 GAAACCTCATCTTCTACTAAAAATACAAAAATTAGCCAGGCGGTGGTGGCGAGGTGCCTGTAA 2783
Db 6547 GAAACCTCATCTTCTACTAAAAATACAAAAATTAGCCGCGGTGGTGGTGGCGAGGTGCCTGTAA 6606

QY 2784 TCCGAGCTACTTGGGAGGCTGAGGCGATGAGAACTTAAACCTGAGAGCGAGGTGAC 2843
Db 6607 TCCGAGCTACTTGGGAGGCTGAGGCGAGGAGATCGGCTTGAACCTAGAGCGAGGTGAC 6666

QY 2844 AGTGAGCCAGATCGTGCCCACTGCATTTCCAGCTGGGCGACAGAGCAAGACTCTGTCTCA 2903
Db 6667 AGTGAGCGGAGATCGTGCCCACTGCATTTCCAGCTGGGCGAGAGACTCTGTCTCA 6726

QY 2904 AAAAAAAAAAAAAA 2917
Db 6727 AAAAAAAAAAAAAA 6740

RESULT 15

US-09-949-016-14964/c
; Sequence 14964, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14964
; LENGTH: 24221
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(24221)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14964

Query Match 8.2%; Score 239.4; DB 3; Length 24221;
Best Local Similarity 88.0%; Pred. No. 4e-48; Mismatches 0; Indels 1; Gaps 1;
Matches 272; Conservative 0; Mismatches 36; Indels 1; Gaps 1;

QY 2610 AAGAAAGGCTGGGCGGCGAGATGGCTCAGCGCTGTAATCCCGAG-ACCTTTGGGAGGCC 2668
Db 23246 AAGCGTAGGGGTGGGCGGCGATGATGGCTCAGCGCTGTAATCCCGGCACTTTGGGAGGCC 23187

QY 2669 GAGGCGAGCGGATCACCTGAGGTGAGGATTTGAGAAAGAGCTGGCCAAACATGGTGAAC 2728
Db 23186 AAGGCGAGGTGGATCACCTGAGGTGAGGATTTGAGAAAGAGCTGGCCAAACATAGTGAAC 23127

QY 2729 CTCATCTCTACTAAAAATACAAAAATTAGCCAGGCGGTGGTGGCGAGTGCCTGTATCCCA 2788
Db 23126 CCGCTCTCTCTAAAAATACAAAAATTAGCCAGGCGATGGTGGCGAGTGCCTGTATCCCTT 23067

QY 2789 GCTACTTTGGGAGGCTGAGGCGATGAGAACTTAAACCTGAGAGCGAGAGGTACAGTGA 2848
Db 23066 GCTACTTTGGGAGGCTGAGGCGAGGAGATCACTTGAACCGGAGGCGGAGGTACAGTGA 23007

QY 2849 GCCAAGATCGTGCCCACTGCATTTCCAGCGCTGGGCGGACAGAGCAAGACTCTGTCTCAAAAA 2908
Db 23006 GCCAAGATCGGACCACTGCATTTCCAGCGCTGGGCGGACAGAGCAAGACTCTCAAAAA 22947

QY 2909 AAAAAAAAAA 2917
Db 22946 AAAAAAGAA 22938

Search completed: March 5, 2006, 04:34:12
Job time : 506 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 03:44:45 ; Search time 2058 Seconds
(without alignments)
11724.998 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gccgagctctgccagccct.....tctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2918	100.0	2918	7	US-10-699-941-3
2	2916.4	99.9	2918	7	US-10-699-941-3
3	2598.2	89.0	2759	6	US-10-104-047-997
4	1438.2	49.3	47219	7	US-10-699-941-10
5	1438.2	49.3	59884	7	US-10-699-941-11
6	1325.8	45.4	2625	6	US-10-104-047-986
7	854.6	29.3	2166	7	US-10-699-941-1
8	675	23.1	703	9	US-10-450-763-16127
9	420.4	14.4	485	3	US-09-918-995-2340
10	330.4	11.3	5656	8	US-10-723-860-4072
11	330.4	11.3	5656	9	US-10-756-149-4072
12	327.6	11.2	11690	9	US-10-487-593-2
13	318.2	10.9	5654	3	US-09-873-367C-644
14	318.2	10.9	5654	3	US-09-968-007A-349
15	318.2	10.9	5654	3	US-09-968-007A-729
16	318.2	10.9	5654	9	US-10-843-641A-644
17	318.2	10.9	5654	9	US-10-843-641A-6819
18	318.2	10.9	5654	9	US-10-843-641A-7199
19	315.8	10.8	2347	6	US-10-220-891-19
20	302.2	10.4	4378	8	US-10-357-930-22138
21	302.2	10.4	4378	8	US-10-357-930-28000
22	267	9.2	786	5	US-10-027-632-152869
23	267	9.2	786	6	US-10-027-632-152869

Sequence 515679, A
Sequence 13015, A
Sequence 34, Appl
Sequence 5729, Ap
Sequence 594866, A
Sequence 263139, A
Sequence 263139, A
Sequence 108, App
Sequence 594865, A
Sequence 55092, A
Sequence 8024, Ap
Sequence 482665, A
Sequence 563212, A
Sequence 251984, A
Sequence 251984, A
Sequence 5691, Ap
Sequence 621024, A
Sequence 84578, A
Sequence 456, App
Sequence 664, App
Sequence 723891, A

US-09-925-065A-515679 623 4
US-09-814-353-13015 660 3
US-10-034-650-34 96595 6
US-10-741-601-5729 403035 7
US-09-925-065A-594866 560 4
US-10-027-632-263139 548 5
US-10-027-632-263139 548 6
US-10-417-375-108 100608 8
US-09-925-065A-594865 560 4
US-09-925-065A-55092 628 4
US-09-764-891-8024 32249 3
US-09-925-065A-482665 562 4
US-09-925-065A-563212 611 4
US-10-027-632-251984 1046 5
US-10-027-632-251984 1046 6
US-10-741-601-5691 67126 7
US-10-741-600-17762 67126 8
US-09-925-065A-621024 483 4
US-09-925-065A-84578 1082 4
US-10-322-281-456 301477 7
US-10-087-192-664 59856 5
US-09-925-065A-723891 2649 4

ALIGNMENTS

RESULT 1

US-10-699-941-3
; Sequence 3, Application US/10699941
; Publication No. US20040146900A1
; GENERAL INFORMATION:
; APPLICANT: Burmeister, Margit
; TITLE OF INVENTION: Ataxia Associated Gene and Protein
; FILE REFERENCE: UM-08441
; CURRENT APPLICATION NUMBER: US/10/699,941
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 2918
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-699-941-3

Query Match 100.0%; Score 2918; DB 7; Length 2918;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCGAGCTCTGCCAGCCCTGAGCTGGGAAGAGAGTACTCGGAGGCGGGCGCA 60
Db 1 GCCGAGCTCTGCCAGCCCTGAGCTGGGAAGAGAGTACTCGGAGGCGGGCGCA 60
Qy 61 GCGGGGCGGCGATGAGAGGGGGCGCAGCCGAGCCCGCTGGGGAGCCACCGCTAAC 120
Db 61 GCGGGGCGGCGATGAGAGGGGGCGCAGCCGAGCCCGCTGGGGAGCCACCGCTAAC 120
Qy 121 CCTGCAACCCACCCACCCCTGCAAAAAGAGCTGGGGGCGCTGGCCACGTCGCTGGG 180
Db 121 CCTGCAACCCACCCACCCCTGCAAAAAGAGCTGGGGGCGCTGGCCACGTCGCTGGG 180
Qy 181 TGACCTTCCTCGGATGAGATCGCCCTCGGAGCATCTCTTCCTCTAGGCTCTGAA 240
Db 181 TGACCTTCCTCGGATGAGATCGCCCTCGGAGCATCTCTTCCTCTAGGCTCTGAA 240
Qy 241 GCGCGGGGAGCGTGGAGGATGCCAGCTGCACCCGGGCGGCTCGCTTTGTTGCA 300
Db 241 GCGCGGGGAGCGTGGAGGATGCCAGCTGCACCCGGGCGGCTCGCTTTGTTGCA 300
Qy 301 GTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCTATCCCTGTCTTAAGCAGTGCCTC 360
Db 301 GTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGGTCTATCCCTGTCTTAAGCAGTGCCTC 360

Db 2521 GAAGACAGACTTTTCTAATGTGGTCCAAATGCGGATCACTGGTCAGATGGAATCTAGAAAG 2580
QY 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAAGGCTGGGCCAGGCAAGATGCTCA 2640
Db 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAAGGCTGGGCCAGGCAAGATGCTCA 2640
QY 2641 CGCCTGTATCCAGACTTTGGGAGGCCGAGGCGGATCACTGAGGTGAGAGTTT 2700
Db 2641 CGCCTGTATCCAGACTTTGGGAGGCCGAGGCGGATCACTGAGGTGAGAGTTT 2700
QY 2701 GAGAACAGCTGGCCAAACATGTGAAACCTCTACTTAAATAATACAAAATTTAGCCA 2760
Db 2701 GAGAACAGCTGGCCAAACATGTGAAACCTCTACTTAAATAATACAAAATTTAGCCA 2760
QY 2761 GCGGTGGTGGCAGGTGCTGTATATCCAGCTACTTTGGGAGGCTGAGGCAATGAGATCACT 2820
Db 2761 GCGGTGGTGGCAGGTGCTGTATATCCAGCTACTTTGGGAGGCTGAGGCAATGAGATCACT 2820
QY 2821 TAAACCTGAGAGCGAGAGTTACAGTGAGCCAAAGATCGTGCCACTGCAATTCAGCCTGG 2880
Db 2821 TAAACCTGAGAGCGAGAGTTACAGTGAGCCAAAGATCGTGCCACTGCAATTCAGCCTGG 2880
QY 2881 CGACAGAGCAAGCTCTGTCTCAAAAAAAGGAAAAA 2918
Db 2881 CGACAGAGCAAGCTCTGTCTCAAAAAAAGGAAAAA 2918

RESULT 2

US-10-699-941-8

; Sequence 8, Application US/10699941

; Publication No. US20040146900A1

; GENERAL INFORMATION:

; APPLICANT: Burmeister, Margit

; TITLE OF INVENTION: Ataxia Associated Gene and Protein

; FILE REFERENCE: UM-08441

; CURRENT APPLICATION NUMBER: US/10/699,941

; CURRENT FILING DATE: 2003-11-03

; NUMBER OF SEQ ID NOS: 62

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 8

; LENGTH: 2918

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-699-941-8

Query Match 99.9%; Score 2916.4; DB 7; Length 2918;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2917; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCGAGCCTCTGCCAGCCTCTGAGCTGGGAAGAGCAGCTACTCGGAGGCGGCGCA 60
Db 1 GCCGAGCCTCTGCCAGCCTCTGAGCTGGGAGNAGAGAGTACTCGGAGGCGGCGCA 60
QY 61 GCGGCGCGGAGTGAAGAGGCGGCGAGCGGAGCCCGCTGGGAGCCCAACGCTAAC 120
Db 61 GCGGCGCGGAGTGAAGAGGCGGCGAGCGGAGCCCGCTGGGAGCCCAACGCTAAC 120
QY 121 CCTGCACCCCAACCCCTGCACAAAGAGCTGGGCGGCGCTGGCAGCTGCGCCTGG 180
Db 121 CCTGCACCCCAACCCCTGCACAAAGAGCTGGGCGGCGCTGGCAGCTGCGCCTGG 180
QY 181 TGACCTTCTCGGATGCAGAAATCCGCCCTCGGAGCATCTCTTCTCTAGGCTCTGAA 240
Db 181 TGACCTTCTCGGATGCAGAAATCCGCCCTCGGAGCATCTCTTCTCTAGGCTCTGAA 240
QY 241 GCGCGGCGGAGGTGAGCGATGCCAGCTGCAACCGGCGAGGCTCGCCTTTGTTGCCA 300
Db 241 GCGCGGCGGAGGTGAGCGATGCCAGCTGCAACCGGCGAGGCTCGCCTTTGTTGCCA 300
QY 301 GTAAGGAGGAGGCTGTCTCAGCTGAGAGGGGTCTCCCTGCTTCAAGCAGTGCCTC 360
Db 301 GTAAGGAGGAGGCTGTCTCAGCTGAGAGGGGTCTCCCTGCTTCAAGCAGTGCCTC 360

```
Db 1441 AACAGCTCTGCTCTGCTCTCAGAGATCAGGAAACAGCATGTCTGAGCGCAGTGAGC 1500
Qy 1501 ATAACAAAGGACATGGAAGAAGATTCAGATGCCAGAAACCTCTGTGAGAGCCCACTG 1560
Db 1501 ATAACAAAGGACATGGAAGAAGATTCAGATGCCAGAAACCTCTGTGAGAGCCCACTG 1560
Qy 1561 GCGCCAGATCTCATCTGCTCTCATCTGAGTCCCAATCTTCAAGGGTCCAGCCCTCC 1620
Db 1561 GCGCCAGATCTCATCTGCTCTCATCTGAGTCCCAATCTTCAAGGGTCCAGCCCTCC 1620
Qy 1621 GTTCATCTCTGAACCCAGCATCTTTTCAGCTGCTTGAACAATGTATTTTTTTTTT 1680
Db 1621 GTTCATCTCTGAACCCAGCATCTTTTCAGCTGCTTGAACAATGTATTTTTTTTTT 1680
Qy 1681 TAAAGATGAGTATTTGTGCTTTCAGAAAAGGGCCAGCTCTGAGCCCTCACCTTCC 1740
Db 1681 TAAAGATGAGTATTTGTGCTTTCAGAAAAGGGCCAGCTCTGAGCCCTCACCTTCC 1740
Qy 1741 ACATCAGAACTCTCAGCCGAGGAAGCAAGACGCAAGGGGTGGCCGCTGCGTC 1800
Db 1741 ACATCAGAACTCTCAGCCGAGGAAGCAAGACGCAAGGGGTGGCCGCTGCGTC 1800
Qy 1801 GGTGGCTCCGCTCTGCTCGAGCCCTGTGCTGAGCTGGATACAGATTCAAGACC 1860
Db 1801 GGTGGCTCCGCTCTGCTCGAGCCCTGTGCTGAGCTGGATACAGATTCAAGACC 1860
Qy 1861 CTTCTCTGCTTGTACCCGCTTCCAGGTTGGAGCCACAGACACCCACCCCGCT 1920
Db 1861 CTTCTCTGCTTGTACCCGCTTCCAGGTTGGAGCCACAGACACCCACCCCGCT 1920
Qy 1921 GGGTCTGCTCTTCTGCTGCTTCTCAGAAATCGGCTCAGACCTAGAACTCA 1980
Db 1921 GGGTCTGCTCTTCTGCTGCTTCTCAGAAATCGGCTCAGACCTAGAACTCA 1980
Qy 1981 ACCCCCTATGAGGCGCAGTCTGGGTAGCTCTGACCTCGACCTTATGTCAAATT 2040
Db 1981 ACCCCCTATGAGGCGCAGTCTGGGTAGCTCTGACCTCGACCTTATGTCAAATT 2040
Qy 2041 TCACACCCATGTTTTCATTTGACCCGCCCTCTCTCATATGACACCCAGCTCC 2100
Db 2041 TCACACCCATGTTTTCATTTGACCCGCCCTCTCTCATATGACACCCAGCTCC 2100
Qy 2101 TTTGAGAGATCAGAGCCCATTTGCAACAAGAGCGCTGCAACCATCTTGTCTCG 2160
Db 2101 TTTGAGAGATCAGAGCCCATTTGCAACAAGAGCGCTGCAACCATCTTGTCTCG 2160
Qy 2161 ATTGCAAAATGACCCCAAGTAACTAGAACATTTCTAGCCCTTTAACTCAGATGCA 2220
Db 2161 ATTGCAAAATGACCCCAAGTAACTAGAACATTTCTAGCCCTTTAACTCAGATGCA 2220
Qy 2221 AGCCACCGGCAAAACCCCGTCAATCTCCACCAAGGAATGAGATATGTGACCTCACT 2280
Db 2221 AGCCACCGGCAAAACCCCGTCAATCTCCACCAAGGAATGAGATATGTGACCTCACT 2280
Qy 2281 GCTCCCCCAACCCAGCGTCAGGCTGGGACACGSCCAACGCTGTTGCGGTTGGAACAGCAG 2340
Db 2281 GCTCCCCCAACCCAGCGTCAGGCTGGGACACGSCCAACGCTGTTGCGGTTGGAACAGCAG 2340
Qy 2341 AGGCTCAGAAATCTGGCTCTGAATAGGACAGACTAGCAGAGGAGATACAGGATATCGG 2400
Db 2341 AGGCTCAGAAATCTGGCTCTGAATAGGACAGACTAGCAGAGGAGATACAGGATATCGG 2400
Qy 2401 GGGTTGAGTGTTCAGAAAGTCAATTCGGAAGATTAATCCAGTGCCTGCGCCGAGCCAC 2460
Db 2401 GGGTTGAGTGTTCAGAAAGTCAATTCGGAAGATTAATCCAGTGCCTGCGCCGAGCCAC 2460
Qy 2461 CTGCATTTCAAGCTTGGACAGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA 2520
Db 2461 CTGCATTTCAAGCTTGGACAGCGGTTCTTGTTCGGGAGGCAAAATTTCCCTAGGAAAA 2520
Qy 2521 GAAGACAGACTTTCTTAATGTGTTCCAAATGCGGATCACTGTGATGATCTTAGAG 2580
```

```
Db 2521 GAAGACAGACTTTCTTAATGTGTTCCAAATGCGGATCACTGTGATGAGTCTTAGAG 2580
Qy 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAAGGCTGGGCCAGGACCGATGCTCA 2640
Db 2581 CACTGAGCTCCCTGTCTCTGGAAGTATTTAAGAAAAGGCTGGGCCAGGACCGATGCTCA 2640
Qy 2641 GCGCTGTAATCCAGACTTTGGAGGCGAGGCGGATCACTCTACTATAAAATACAAAATTAGCCA 2700
Db 2641 GCGCTGTAATCCAGACTTTGGAGGCGAGGCGGATCACTCTACTATAAAATACAAAATTAGCCA 2700
Qy 2701 GAGAACAGCTGGGCCCAACATGTTGAACTCATCTCTACTATAAAATACAAAATTAGCCA 2760
Db 2701 GAGAACAGCTGGGCCCAACATGTTGAACTCATCTCTACTATAAAATACAAAATTAGCCA 2760
Qy 2761 GCGCTGTTGAGGAGGCTGTAATCCAGCTTACTTTGGAGGCTGAGGATGAGAACTACT 2820
Db 2761 GCGCTGTTGAGGAGGCTGTAATCCAGCTTACTTTGGAGGCTGAGGATGAGAACTACT 2820
Qy 2821 TAAACCTGAGAGGCGAGGTTTACAGTGAAGCAAGATCGTCCACTGCAATCCAGCTGG 2880
Db 2821 TAAACCTGAGAGGCGAGGTTTACAGTGAAGCAAGATCGTCCACTGCAATCCAGCTGG 2880
Qy 2881 CGACAGCAAGACTCTCTCTCAAAAAA 2918
Db 2881 CGACAGCAAGACTCTCTCTCAAAAAA 2918

RESULT 3
US-10-104-047-997
; Sequence 997, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 997
; LENGTH: 2759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-997

Query Match 89.0%; Score 2598.2; DB 6; Length 2759;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 2756; Conservative 0; Mismatches 3; Indels 136; Gaps 2;

Qy 8 CTCTGCCAGCCCTGAGCTGGGAAGAGCAGCTACTCTGGAGGAGGCGCGGCGCGGCG 67
Db 1 CTCTGCCAGCCCTGAGCTGGGAAGAGCAGCTACTCTGGAGGAGGCGCGGCGCGGCG 60
Qy 68 GCGCATGAGAGGCGGCGCAGCGCGAGCCCGCGCTGGGAGGCCACCGCTAACCTTGCAC 127
Db 61 GCGCATGAGAGGCGGCGCAGCGCGAGCCCGCGCTGGGAGGCCACCGCTAACCTTGCAC 120
Qy 128 CCCACCCACCCCTGCAAAAAGAGCTGGCGGCGCTGGGCCACGTCGCTTGGGTGACCTT 187
Db 121 CCCACCCA----- 128
Qy 188 CCTCGATGCAAGATCCCGCTCGAGCATCTCTTCTCTCTAGGCTCTGAGGCCCGG 247
Db 129 -----CCCTTGCAGCATCTCTTCTCTCTAGGCTCTGAGGCCCGG 170
Qy 248 GAGCGCTGAGCGATGCCAGCTGCAACCGGCGAGGCTCGCTTGTGTTGTCAGTAAGA 307
Db 171 GAGCGCTGAGCGATGCCAGCTGCAACCGGCGAGGCTCGCTTGTGTTGTCAGTAAGA 230
Qy 308 GAGAGGCTGTCTGAGCTGCAAGAGGCTCATCTCTGCTTCAAGCAGTGCCTTCCAG 367
```

Db	231	GGAGAGGCTGTCTCAGCTGCAGAGGGGTATCCCTGCTTCAAGCCAGTGCCTCTTCCAG	290
Qy	368	CTCCCATGGGGAACAACGAGGCAACGCTCCGATGAGAAACGTTGACGTTGAAGAGAAAT	427
Db	291	CTCCCATGGGGAACAACGAGGCAACGCTCCGATGAGAAACGTTGACGTTGAAGAGAAAT	350
Qy	428	GGCAGGACGAGATCTTCCAGGCCACTCCAGAGAGAGCGGGGTGGAACTGCTTGGCA	487
Db	351	GGC-----	353
Qy	488	GCCGGTGGGAAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCGCATCGTA	547
Db	354	-----AAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCGCATCGTA	404
Qy	548	AGAGGAGACGCTGGTGGGCCCAAGAGATCAACATTTCTCTGGATCAGAGTGGGGTCCC	607
Db	405	AGAGGAGACGCTGGTGGGCCCAAGAGATCAACATTTCTCTGGATCAGAGTGGGGTCCC	464
Qy	608	TGCTGTCCGATGACTTCTTGGATACCCCTGATGATGATGATGATGATGATGATGATGATG	667
Db	465	TGCTGTCCGATGACTTCTTGGATACCCCTGATGATGATGATGATGATGATGATGATGATG	524
Qy	668	AGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAAATGCGAACGAACTGGAGTGG	727
Db	525	AGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAAATGCGAACGAACTGGAGTGG	584
Qy	728	AAGACGACACCCCGTGGGCAACCGCCAAAGAAACATATGCGGGGACAGCGGGATCTATTG	787
Db	585	AAGACGACACCCCGTGGGCAACCGCCAAAGAAACATATGCGGGGACAGCGGGATCTATTG	644
Qy	788	GGGACGGACGACGAGGAGCGGAGCGCGCCGCAACGCGGCGCTGTGGCGGACAGTGAATCA	847
Db	645	GGGACGGACGACGAGGAGCGGAGCGCGCCGCAACGCGGCGCTGTGGCGGACAGTGAATCA	704
Qy	848	TGCGGAGCAGAGACCGCTATAGACTGCAATGATCCGGCTTACATGAAGTGTCA	907
Db	705	TGCGGAGCAGAGACCGCTATAGACTGCAATGATCCGGCTTACATGAAGTGTCA	764
Qy	908	CCACGGAGGTTACTAGCGGAGGCTTCAAGCCATCATCGTCTTTCGAGCGCTGCTCC	967
Db	765	CCACGGAGGTTACTAGCGGAGGCTTCAAGCCATCATCGTCTTTCGAGCGCTGCTCC	824
Qy	968	TTCCAGACGAGCGCTCCCGATACCACTATCATGAGGAACTCTTCTGTGATGCTCA	1027
Db	825	TTCCAGACGAGCGCTCCCGATACCACTATCATGAGGAACTCTTCTGTGATGCTCA	884
Qy	1028	TCAGCAGCTTACAGCTCCTGGTGGCTGAGGACTACATGATCGTGTACCTGAGCGGTGCA	1087
Db	885	TCAGCAGCTTACAGCTCCTGGTGGCTGAGGACTACATGATCGTGTACCTGAGCGGTGCA	944
Qy	1088	CGCCCGCGGAGGATCGCTGGAAATCGGCTGGAAGAGTGCTACCAAGATGATCGAC	1147
Db	945	CGCCCGCGGAGGATCGCTGGAAATCGGCTGGAAGAGTGCTACCAAGATGATCGAC	1004
Qy	1148	GGAGTTGGGAAAAAAGCTGAGTCTTGATCATCGTCCACCCCTCGTGGTTCAATTCGA	1207
Db	1005	GGAGTTGGGAAAAAAGCTGAGTCTTGATCATCGTCCACCCCTCGTGGTTCAATTCGA	1064
Qy	1208	CTGTGTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCAATCAACAGATCCAGTAG	1267
Db	1065	CTGTGTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCAATCAACAGATCCAGTAG	1124
Qy	1268	TGCACAGCTTGGAGACCTGGAGCAACTCATCCCTATGGAACACGTCAGATCCAGACT	1327
Db	1125	TGCACAGCTTGGAGACCTGGAGCAACTCATCCCTATGGAACACGTCAGATCCAGACT	1184
Qy	1328	GGGTCTGCAATACGAGAGGAAAGACTGGAAGCCAGGAGGAGCGCGGCCCCAGC	1387
Db	1185	GGGTCTGCAATACGAGAGGAAAGACTGGAAGCCAGGAGGAGCGCGGCCCCAGC	1244
Qy	1388	CGGAGTTGTGCTGCCAGGTCTGAAGAGAGGAGGAGTGGCAACAGTGGAAAAACAGT	1447
Db	1245	CGGAGTTGTGCTGCCAGGTCTGAAGAGAGGAGGAGTGGCAACAGTGGAAAAACAGT	1304
Qy	1448	CTGCTCTGCTCTCAGAGATCAGGAAACAGACATGTCTCTAGGCGCAGCGAGATAACAA	1507
Db	1305	CTGCTCTGCTCTCAGAGATCAGGAAACAGACATGTCTCTAGGCGCAGCGAGATAACAA	1364
Qy	1508	AGGACATGAGAGAGATTCAGATGCCAGAAACCTCTGTGTGACAGCCACCTGGCCCCAG	1567
Db	1365	AGGACATGAGAGAGATTCAGATGCCAGAAACCTCTGTGTGACAGCCACCTGGCCCCAG	1424
Qy	1568	ATCTCATCTGCTCTCATCTGAGTCCCAATCTTCAAGGGGTGCGAGCCCTCGGTTATC	1627
Db	1425	ATCTCATCTGCTCTCATCTGAGTCCCAATCTTCAAGGGGTGCGAGCCCTCGGTTATC	1484
Qy	1628	TCTGAAACCCAGCATCTCTTTCAGCTGTCTGAAACATTTATTTTTTTTTTTTAAACGAT	1687
Db	1485	TCTGAAACCCAGCATCTCTTTCAGCTGTCTGAAACATTTATTTTTTTTTTTTAAACGAT	1544
Qy	1688	GCAGTATTTGTGCTTCAGAAAGGGCCAGCTCTGAGCCCTCACCTTCCACATCA	1747
Db	1545	GCAGTATTTGTGCTTCAGAAAGGGCCAGCTCTGAGCCCTCACCTTCCACATCA	1604
Qy	1748	CGAACTCTCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1807
Db	1605	CGAACTCTCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1664
Qy	1808	TCCGCTCTCTGCTCGCAGCCCTGTGTGCTCAGAGCTGATATCAAGATTTCAAGACCTTCTCT	1867
Db	1665	TCCGCTCTCTGCTCGCAGCCCTGTGTGCTCAGAGCTGATATCAAGATTTCAAGACCTTCTCT	1724
Qy	1868	TGCTTGTCAACCGCTTCCAGGTTGGAGCCACAGACCCACCGCCACCCCGGCTGGGCTG	1927
Db	1725	TGCTTGTCAACCGCTTCCAGGTTGGAGCCACAGACCCACCGCCACCCCGGCTGGGCTG	1784
Qy	1928	CGTCTCTCTGCTGCTTCCCTCCAGAAATGCGGCTCAGACCTCAGAAAGCTCAACCCGCC	1987
Db	1785	CGTCTCTCTGCTGCTTCCCTCCAGAAATGCGGCTCAGACCTCAGAAAGCTCAACCCGCC	1844
Qy	1988	TATGAGGCGCAGCTCTCGGGTAGCTCTGAGCTCCGACCTTATGTCCAAATTTCCACACC	2047
Db	1845	TATGAGGCGCAGCTCTCGGGTAGCTCTGAGCTCCGACCTTATGTCCAAATTTCCACACC	1904
Qy	2048	CATGTTTTTTCATTTGACCCCGCTTCTCGCTCATATATGACACCCAGCTCTTTTGAGA	2107
Db	1905	CATGTTTTTTCATTTGACCCCGCTTCTCGCTCATATATGACACCCAGCTCTTTTGAGA	1964
Qy	2108	GGATCAGAGCCATTTGCAAGAGAGCGCTGCCAACCATCTTGTCTCTCGATGCA	2167
Db	1965	GGATCAGAGCCATTTGCAAGAGAGCGCTGCCAACCATCTTGTCTCTCGATGCA	2024
Qy	2168	AATGACACCCAGTAATCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGTCAGGCCACC	2227
Db	2025	AATGACACCCAGTAATCTAGAACATTTCTCAAGCCCTTTTAACTCAGATGTCAGGCCACC	2084
Qy	2228	GGGCAAAACCCGCTCAATCTCCCAACAGGAATGAGATATGTGACCTCACTGCTCCCC	2287
Db	2085	GGGCAAAACCCGCTCAATCTCCCAACAGGAATGAGATATGTGACCTCACTGCTCCCC	2144
Qy	2288	CAACCCAGCGTACGCTGGGACACGCGCAACGCTGTTCCGGGTTGGAAACAGAGAGGCTCA	2347
Db	2145	CAACCCAGCGTACGCTGGGACACGCGCAACGCTGTTCCGGGTTGGAAACAGAGAGGCTCA	2204
Qy	2348	GAAACTGCTCTGAAATAGGACACCTAGCAAGAGGAGATACAGGGTATCGGGGTTG	2407
Db	2205	GAAACTGCTCTGAAATAGGACACCTAGCAAGAGGAGATACAGGGTATCGGGGTTG	2264
Qy	2408	AGTGTCTCAGAACTCATTTCCGGAAAGATAAATTCAGTGGCTGGCGCAGCACCTGCAAT	2467
Db	2265	AGTGTCTCAGAACTCATTTCCGGAAAGATAAATTCAGTGGCTGGCGCAGCACCTGCAAT	2324
Qy	2468	CAAAAGTTGACAGCGGTTCTTGTTCGGAGGCAAAATTTCCCTTAGGAAAAAGAGACA	2527
Db	2325	CAAAAGTTGACAGCGGTTCTTGTTCGGAGGCAAAATTTCCCTTAGGAAAAAGAGACA	2384

QY 305 GGAGGAGAGCTCTCTCAGCTGACAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTCTTCC 364
DB 241 GGAGGAGAGCTCTCTCAGCTGACAGGGGTCACTCCCTGCTTCAAGCCAGTGCCTCTTCC 300
QY 365 CAGCTCCATGGGGACCAAGAGCCAGCTCCGGATGGAAACGCTGGAGCGTGAAGAGG 424
DB 301 CAGCTCCATGGGGACCAAGAGCCAGCTCCGGATGGAAACGCTGGAGCGTGAAGAGG 360
QY 425 AATGGCAGGACGAAGATCTTCCAGAGCCACTCCCAAGAGAGCGGGGTGGAACTGCTTG 484
DB 361 AATGGCAGGACTAAGATCTTCCAGGACCACTCCCAAGAGAGCGGGGTGGAACTGCTTG 420
QY 485 GCAGCCCGGTGGAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCGCATC 544
DB 421 GCAGCCCGGTGGAGACACATCTCTCTCCCAACACGCTAAATTTCAACGGAGCGCATC 480
QY 545 GTAAGAGGAGAGCGTGGTGGCCCAAGAGATCAACATTTCTCTGGATCAGAGTGGGGGT 604
DB 481 GCAAGAGGAGAGCGTGGTGGCCCAAGAGATCAACATTTCTCTGGATCAGAGTGGGGGT 540
QY 605 CCTGCTGCTCCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACA 664
DB 541 CCTGCTGCTCGATGACTTCTTGGATACCCCTGATGACCTGGATATTAACGTGGATGACA 600
QY 665 TCGAGACCCCGGTGAGACCGACTCGCTGGAGTTCTGGGGATGGCAACGAACTGGAGT 724
DB 601 TCGAGACCCCGGTGAGACCGACTCGCTGGAGTTCTGGGGATGGCAACGAACTGGAGT 660
QY 725 GGGAGACGACACCCCGGTGGCCACCGCCAGAACATGCCCCG 767
DB 661 GGGAGACGACACCCCGGTGGCCACCGCCAGAACATGCCCCG 703

RESULT 9

US-09-918-995-2540
; Sequence 2540, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2540
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(485)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2540

Query Match 14.4%; Score 420.4; DB 3; Length 485;
Best Local Similarity 99.8%; Pred. No. 8.2e-111;
Matches 421; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 346 TCAAGCAGTGCTCTTCCAGCTCCCATGGGACCAACAGCCAGCGTCCGGATGGAA 405
DB 64 TTAAGCAGTGCTCTTCCAGCTCCCATGGGACCAACAGCCAGCGTCCGGATGGAA 123
QY 406 AACGTGGACGTGAAGGAGGAATGGCAGGACGAGATCTTCCAGGGCACTCCCAAGAG 465
DB 124 AACGTGGACGTGAAGGAGGAATGGCAGGACGAGATCTTCCAGGGCACTCCCAAGAG 183
QY 466 ACGGGGTGAACTGCTTGGCAGCCCGGTGGAAAGACATCTCTCTCCCAACAGCTA 525

DB 184 ACGGGGTGGAACACTCTTGGCAGCCGGTGGAAAGACACATCTCTCTCCCAACAGCTA 243
QY 526 AATTTCAACGGAGCCGATCGTAAGAGGAGAGCGTGGTGGCCCAAGAGATCAACATTTCT 585
DB 244 AATTTCAACGGAGCCGATCGTAAGAGGAGAGCGTGGTGGCCCAAGAGATCAACATTTCT 303
QY 586 CTGGATCAGAGTGAAGGGTCCCTGCTGCTCCGATGACTTCTTGGATACCCCTGATGACCTG 645
DB 304 CTGGATCAGAGTGAAGGGTCCCTGCTGCTCCGATGACTTCTTGGATACCCCTGATGACCTG 363
QY 646 GATATTAACGTGGATGACATCGAGACCCCGGATGAGACCGACTCGCTGGAGTTCTGGGG 705
DB 364 GATATTAACGTGGATGACATCGAGACCCCGGATGAGACCGACTCGCTGGAGTTCTGGGG 423
QY 706 AATGCAACGAGTGGAGTGGAGACGACACCCCGGTGGCCACCGCCCAAGAACATGCC 765
DB 424 AATGCAACGAGTGGAGTGGAGACGACACCCCGGTGGCCACCGCCCAAGAACATGCC 483
QY 766 GG 767
DB 484 GG 485

RESULT 10

US-10-723-860-4072
; Sequence 4072, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NFUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4072
; LENGTH: 5656
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-4072

Query Match 11.3%; Score 330.4; DB 8; Length 5656;
Best Local Similarity 63.3%; Pred. No. 2.5e-84;
Matches 558; Conservative 0; Mismatches 291; Indels 33; Gaps 2;
QY 469 GGGGTGGAACCTGCTTGGCAGCCCGGTGGAGACACATCTCTCTCCCAACAGCTAAT 528
DB 1510 GAGGCGGTGAGTCCGAGTCCGAGATGAGGCTGACAGATGAGGCTCTCTCTCTGAT 1569
QY 529 TTCAACGGAGCGCATCGTAAGAGGAGAGCGTGGTGGCCCAAGAGATCAACATTTCTCTG 588
DB 1570 CTTAATGACATCTCTCTCGGAGATCAAGCTCACAGCCCCCAATATCAATCTTCTCTG 1629
QY 589 GATCAGAGTGAAGGGTCCCTGCTGCTCCGATGACTTCTTGGATACCCCTGATGACCTGGAT 648
DB 1630 GACCAAGAGTGAAGGATCTATTCTCTGATGATACTTGGACAGTCCAGATGAAATGAC 1689
QY 649 ATTAACGTGGATGACATCGAGACCCCGGATGAGACCGACTCGCTGGAGTTCTTGGGAT 708
DB 1690 ATCAATGTGGATGAACCTTGATACCCCGGATGAAAGAGATTTCTTTGATACACTGGCCAT 1749
QY 709 GSCAACGAACTGGAGTGGGAGAGACGACACCCCGGTGGCCACCGCCCAAGAACATGCCCCGG 768
DB 1750 G-----AAGATCCCAAGCCCAACAAAGATTTCTGGC 1779
QY 769 GACACCGCGGATCTATTGTTGGGAGCGGACGAGAGGAGCGGAGCGCCGCGCCAAACGGGGC 828
DB 1780 CAAGAGTCAGAGTCTATTTCAGAAATATACGGCCGAGAGGAGGAGGAGCAAC---CGG 1836

QY 829 CTGTGGCGGACAGTATCGGGAGCAAGAGCACCGTATAGACCTGCAATGATCCGG 888
Db 1837 CTTTGGAGGACAGTGTCTATGGAGAACAGAGCAGCGCATTCACATGAAGGTCAATCGAG 1896
QY 889 CCTTACATGAAGTGGTCAACCGAGGAGTACTACGGCAAGCCCTCAAGCCCATCATC 948
Db 1897 CCTACAGGAGAGTCAATTCACGGAGGATCTATGGGACCGGTCTAAATGCCATCAT 1956
QY 949 GTCTTCGACGCTGCTTCTCCAGACAGCAGCAGCTCCCGACTACCACTACATCATGGAG 1008
Db 1957 GTGTTGGCGGCTGTTTCTGCCAGACAGCAGTGGGCGGATTCACATATGTCTATGGAA 2016
QY 1009 AACCTCTTCTGTACGTATCAGAGCTTGAAGCTCTGTGTCTGAGGACTACATGATC 1068
Db 2017 AATCTTTTCTATATGTAATAGTACTTTAGAGTTGATGTAGCTGAAGACTATATGAT 2076
QY 1069 GTGTACTGAAGCTGACCGCCCGCGGAGGATGCTTGGATCGGCTGGCTGAAGAG 1128
Db 2077 GTGTACTGAATGCTGCAACCCCAAGAGGAGATGCGAGGGCTAGGCTGGATGAAGAA 2136
QY 1129 TGTACACAGATGACAGCGAGGTGCGGAAACCTGAAGTCTTGTATCATCGTCCAC 1188
Db 2137 TGCTACACAGATGATGACAGCGGTGAGGAGAGATTTGAATCATTCATCTGTTTAT 2196
QY 1189 CCTCTGTGTCTATCGGACTGTGTGGCCATCTCTGCGCTTTCATCAGCGTCAAGTTC 1248
Db 2197 CCATCTTGGTTTCTCAGCAATCTTGTGTGACAGACCTTTTATAGTTCAAAATTC 2256
QY 1249 ATCAACAGATCAGTACGTGACAGCTGGAGAGCTGGAGCAACTCATCTCCATGGAA 1308
Db 2257 AGCAGTAAATTAATATGTCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGAT 2316
QY 1309 CAGTCCAGATCCAGACTGCGTCTGCAATACGAGAGGAA 1350
Db 2317 TGCATCCACTCCAGAGGATCATCAAACTGGATGAAGAA 2358

RESULT 11

US-10-756-149-4047
; Sequence 4047, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4047
; LENGTH: 5656
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-4047

Query Match 11.3%; Score 330.4; DB 9; Length 5656;
Best Local Similarity 63.3%; Pred. No. 2.5e-84;
Matches 558; Conservative 0; Mismatches 291; Indels 33; Gaps 2;

QY 469 GGGGTGGAAGTCTTGGAGCCGGTGGAGACACATCTCTCCCTCCCAACAGCTAAT 528
Db 1510 GAGGGGTCTGAGTCCCAAGTCTGAGACATGAGCGCTGAACCTCTTAATCTCTGGAT 1569
QY 529 TTCAACGGAGCGCATCGTGAAGAGAGAGCGCTGGTGGCCCGCCAGAGATCAAACTTTCTGT 588
Db 1570 CTTAATGACACTCATCTCGGAGATCAAGCTCAGAGCCCAATATCATCTTTCTGT 1629
QY 589 GATCAGAGTGAAGGGTCCCTGCTGCCGATGACTTTCTGGATPACCCCTGATGACCTGGAT 648
Db 1630 GACCAAGAGTGAAGGATCTATCTCTCTGATGATACTTGGACAGTCCAGATGAAATTCAG 1689

QY 649 ATTAACTGTGATGACATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCCTGGGAAT 708
Db 1690 ATCAATGTGTGATGACTTGTATCCCGGATGAGACGATTTCTTTGAGTACACTGGCCAT 1749
QY 709 GGCACGAATCTGGAGTGGGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCCGG 768
Db 1750 G-----AAGATCCCAAGCCCAAGATTTGGC 1779
QY 769 GACAGCGCGATCTATTTGGGACGGCAGCAGAGGAGCGGAGCGCCGCCCAACGGGCGC 828
Db 1780 CAAGAGTCAAGTCTATTTCCAGATATATACGGCCCAAGAGAGAAACGGGAGGACAAAC---CG 1836
QY 829 CTGTGCGGACAGTGTATCATCTCGGGAGCAAGAGCAGCTATAGACTGTCACATCATGATCCCG 888
Db 1837 CTTTGGAGGACAGTGTCTATTTGAGAACAGAGAGCGGATTTGACATGAAGGTCAATCGAG 1896
QY 889 CTTTACATGAAGTGTCTACCCAGCGAGGTTACTAGCGGAAAGGCTCAACGCGCATCATC 948
Db 1897 CCTACAGGAGTCAATTTCTCAGCGAGGATATATGGGACGGTCTAATGCCATCAT 1956
QY 949 GTCTTGGAGCGCTGCTTCTTCCAGACAGCAGCGCTCCCGACTACCACTATCATGAGAG 1008
Db 1957 GTGTTTGGCGGCTGTTTCTGCGCAGACAGCTCGGGCGGATTTACCACTATGTCATGGAA 2016
QY 1009 AACCTCTTCTGTAGTCTATCAGCAGCTTAGAGCTCTGCTGGCTGAGGACTACATGATC 1068
Db 2017 AATCTTTTCTATATGTAATAGTACTTTAGAGTTGATGGTGTGAGGACTATATGAT 2076
QY 1069 GTGTACTGAAACGGTGCACACCGCGGAGGATGCTTGGAAATCGGCTGGCTGAAGAG 1128
Db 2077 GTGTACTTGAATGTGCAACCCCAAGAGAGGAGATGCGAGGCTAGGCTGGATGAAGAA 2136
QY 1129 TGCTACACAGATGATCGACCGGAGGTGCGGAAAAACCTGAACTCTTGTATCATCTCGTCCAC 1188
Db 2137 TGCTACACAGATGATGACAGCGGTGAGGAGAAATTTGAATCATTCATCTGTTTAT 2196
QY 1189 CCTCTGTGTCTATCGGACTGTGTGGCCATCTCTGCGCTTTCATCAGCGTCAAGTTC 1248
Db 2197 CCATCTTGGTTCATCAGAACTCTTGTGTGACAGCACTTTTATAGTTCAAAATTC 2256
QY 1249 ATCAACAGATCCAGTACGTGACAGCTTGGAGACCTGGAGCAACTCATCTCCATGGAA 1308
Db 2257 AGCAGTAAATTAATATGTCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGAT 2316
QY 1309 CAGTCCAGATCCAGACTGCGTCTGCAATACGAGAGGAA 1350
Db 2317 TGCATCCACTCCAGAGGATCATCAAACTGGATGAAGAA 2358

RESULT 12

US-10-487-593-2
; Sequence 2, Application US/10487593
; Publication No. US20050064413A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWARA, AKIRA
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: NOVEL HUMAN BMCC1 GENE
; FILE REFERENCE: 7388/80977
; CURRENT APPLICATION NUMBER: US/10/487,593
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: PCT/JP02/08520
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: JP 2001-255198
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 2
; LENGTH: 11690
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-487-593-2

Query Match 11.2%; Score 327.6; DB 9; Length 11690;
Best Local Similarity 63.2%; Pred. No. 2.2e-83;
Matches 557; Conservative 0; Mismatches 289; Indels 36; Gaps 2;

QY 469 GGGGTGGAACTGCTTGGCAGCCGGTGGAGACACATCTCTCTCCCAACAGCTAAAT 528
DB 7547 GAGGGGTGTGAGTCCAGTGTGCAGACATGAGGCCCTGAACCTCTAATTTCTCGAT 7606
QY 529 TTCAACGGAGCGCATCGTAAGAGGAAGACGCTGGTGGCCCCAGAGATCAACATTTCTCTG 588
DB 7607 CTTAATGACACTCATCTCGAGAAATCAAGCTCAGAGCCCCAATATCAATCTTCTCTG 7666
QY 589 GATCAGAGTGAAGGGTCCCTGCTGTCGATGACTTTTGGATACCCCTGATGACCTGGAT 648
DB 7667 GACCAAGTGAAGGATCTATTCTCTGTGATGATACTTGGACAGTCCAGATGAATTGAC 7726
QY 649 ATTAAGTGGATCAGATCGAGACCCCGATGACACGCTCGTGGAGTCTCTGGGAAT 708
DB 7727 ATCAATGTGGATGAATCTTGATACCCCGATGAAGCAGATTTCTTTGATACACTGG- 7782
QY 709 GGCAACGAAGTGGTGGGAAGACGACACCCCGTGGCCACCGCCCAAGAACATGCCCGG 768
DB 7783 -----CCATGATCCACAGCAAAAGATTTCTGGC 7813
QY 769 GACAGCGCGGATCTATTTTGGGAGCGGCACGACGAGGAGCGGACGCGCCCAACGGCGC 828
DB 7814 CAAGAGTCAGAGTCTATTCCAGAAATATAGCCGGAAGAGGAACGGGAGGACAA- 7870
QY 829 CTGTGGGAGCAGTGATCATCGGGGAGCAAGACACGCTATAGACCTGACATGATCCGG 888
DB 7871 CTTTGGAGGACAGTGGTCAATTTGAGAAACAGAGCAGCGCATTTGACATGAAGGTCATCGAG 7930
QY 889 CTTTACATGAAGTGGTCAACGAGGAGTACTACGGGAGGAGCGCTCAACGCCATCATC 948
DB 7931 CCCTACAGGAGATCAATTTCTACAGGAGATCTATGGGACGGTCTAAATGCCATCAT 7990
QY 949 GTCTTGGAGCTGTCTTCTTCCAGACAGCAGCGCTCCCGACTACCACTACATCATGAG 1008
DB 7991 GTGTTTGGCGCTGTTTCTGCCAGACAGCAGTCCGGCGGATTTACCACATATGTCATG 8050
QY 1009 AACCTTCTCTGTACGTCATCAGCAGCTTAGAGCTCTGTGGCTGAGAGCTACATGATC 1068
DB 8051 AATCTTTTCTATATGTAATAGTACTTTTAGAGTTGATGTAGTCTGAAGACTATATGAT 8110
QY 1069 GTGTACTGAAAGCTGACCGCCCGCGGAGGATGCTTGAATCGGCTGGCTGAAGAA 1128
DB 8111 GTGTACTGTAATGGTGGAAACCCCAAGAGGAGATGCCAGGGCTAGGCTGATGAAGAA 8170
QY 1129 TGCTACAGATGATCGACCGAGGTTTGGGAAACCTGAAGTCTTGTATCATCGTCCAC 1188
DB 8171 TGCTACAGATGATGACAGACGTTTGGAGAAATTTGAAATCATTTATCATTTGTCAT 8230
QY 1189 CCCTCGTGTTCATTGCGACTGTGCGCATCTCTCGCCCTTTTCATGACGGTCAAGTTC 1248
DB 8231 CCATCTTGGTTTCATCAGAACATCTTGTGTGTGACACGACCTTTTATAAGTTTCAAAATTC 8290
QY 1249 ATCAACAGATGATCAGTACGTCAGCTTGGAGACCTGGAGCACTCATCTCCCTATGGA 1308
DB 8291 AGCAGTAAATTAATATGTAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGAT 8350
QY 1309 CAGTCCAGATCCAGATCGCTGCTGCAATACGAAGAGAA 1350
DB 8351 TGCATCCATTCAGAGACATCATCAAACTGGATGAAGAA 8392

RESULT 13

US-09-873-367C-644
; Sequence 644, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel

APPLICANT: Endress, Gregory
APPLICANT: Augustus, Meena
APPLICANT: Ebner, Reinhard
APPLICANT: Carter, Kenneth
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
FILE REFERENCE: Signature Gene Sets
CURRENT APPLICATION NUMBER: US/09/873,367C
CURRENT FILING DATE: 2003-04-29
PRIOR APPLICATION NUMBER: U.S. 60/236,891
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: U.S. 60/236,842
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: U.S. 60/244,867
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: U.S. 60/245,084
PRIOR FILING DATE: 2000-11-01
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: Patent in version 3.0
SEQ ID NO 644
LENGTH: 5654
TYPE: DNA
ORGANISM: Homo sapiens
US-09-873-367C-644

Query Match 10.9%; Score 318.2; DB 3; Length 5654;
Best Local Similarity 63.2%; Pred. No. 8.7e-81;
Matches 558; Conservative 0; Mismatches 288; Indels 37; Gaps 3;

QY 469 GGGGTGGAACTGCTTGGCAGCCGGTGGAGACACATCTCTCTCCCAACAGCTAAAT 528
DB 1510 GAGGGGTGTGAGTCCAGTGTGCAGACATGAGGCCCTGAACCTCTAATTTCTCGAT 1569
QY 529 TTCAACGGAGCGCATCGTAAGAGGAGAGCGTGGTGGCCCCAGAGATCAACATTTCTCTG 588
DB 1570 CTTAATGACACTCATCTCTCGGAGAAATCAAGCTTCAAGCCCCAATATCAATCTTTCTCTG 1629
QY 589 GATCAGAGTGAAGGGTCTCTGTCTCGATGACTTTCTTGGATACCCCTGATGACCTGGAT 648
DB 1630 GACCAAGTGAAGGATCTATTCTCTGATGATACTTTGGACAGTCCAGATGAATGAC 1689
QY 649 ATTAAGTGGATGATCATCGAGACCCCGATGAGACCGACTGCTGCTGGAGTCTCTGGGAAAT 708
DB 1690 ATCAATGTGGATGAATTTGATACCCCGATGAAGCAGATCTTTTGGAGTACACTGG- 1745
QY 709 GGCAACGAAGTGGAGTGGGAGACACACCCCGTGGCCACCGCCCAAGACATGCCCGG 768
DB 1746 -----CCATGATCCCAAGCAAAAGATTTCTGGC 1776
QY 769 GACAGCGCGGATCTATTTTGGGAGCGGCACGAGGAGCGGACGCGCCCAACGGCGC 828
DB 1777 CAAGAGTCAAGTCTATTTCAGAAATATAGCGCCGAAGAGGAAACGGAGGACAA- 1833
QY 829 CTGTGGGAGCAGTGTATCATCGGGAGCAAGAGACCGCTATAGACTGACATGATATCGG 888
DB 1834 CTTTGGAGGACAGTGGTCAATTCATGGAGAAACAGAGCAGCGCATTTGACATGAAGGTCATCGAG 1893
QY 889 CCTTACATGAAGTGGTCAACCCAGGAGGTTACTAGCGGAGGCTTCAACGCCATCATC 948
DB 1894 CCTTACAGAGAGTCAATTTCTCAGGAGGATATATGGGAGCGGTCTAAATGCCATCAT 1953
QY 949 GTCTTGGAGCTGTCTCTTCCAGACAGCAGCTCCCGACTACCACTATCATCATGAG 1008
DB 1954 GTGTTTGGCGCTGTTTCTGCCAGACAGCAGTCCGGCGGATTAACACTATGTCATGAA 2013
QY 1009 AACCTTCTCTGTATGATCATCAGAGCTTAGAGTCTCTGTGGTGGCTGAGAGCTATCATGATC 1068
DB 2014 AATCTTTTCTATATGTAATAGTACTTTAGAGTGTAGTGTGAGTGAAGACTATATGAT 2073
QY 1069 GTGTACCTGMAAGGTCACCGCCC- CGGCGGAGGATGCTTGAATCGGCTGGCTGAAGAA 1127
DB 2074 GTGTACTTGAATGGTGGCAACCCCGGAGGAGGATGCCAGGGCTAGGCTGGATGAAGAA 2133

QY 1128 GTCTACCATGATGACCGGAGTTGCGGAAACCTGAACTCCTTGATCATCGTCCA 1187
DB 2134 ATGCTACCATGATGACGAGCGTTGAGGAAGAAATTTGAAATCATTTCTTCA 2193
QY 1188 CCCTCTGTGTTCAATCGGACTGTGCTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTT 1247
DB 2194 TCCATCTTGTTTCAATCAGAACATCTTGTGTGACAGGACCTTTTATAGTTCAAAAT 2253
QY 1248 CATCAACAAGATCCAGTACGTGACACAGCTTTGGAAGACCTGGAGCAACTCATCCCTATGGA 1307
DB 2254 CAGCAGTAAATTAATATGCTCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGA 2313
QY 1308 ACAGTCCAGATCCAGATGCGTCTGCTGCAATACGAAGAGGAA 1350
DB 2314 TTGCATCCCAATTTCCAGAGAGCATCATCAAACTGGATGAAGAA 2356

RESULT 14

US-09-968-007A-349

; Sequence 349, Application US/09968007A
; Publication No. US20040115625A1

GENERAL INFORMATION:

; APPLICANT: Ebner, Reinhard

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signa
; FILE REFERENCE: 689290-71

; CURRENT APPLICATION NUMBER: US/09/968, 007A
; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 172
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 173
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 278
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 294
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 295
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 316
; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 1001
; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 349
; LENGTH: 5654

; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-968-007A-349

Query Match 10.9%; Score 318.2; DB 3; Length 5654;
Best Local Similarity 63.2%; Pred. No. 8.7e-81;
Matches 558; Conservative 0; Mismatches 288; Indels 37; Gaps 3;
QY 469 GGGGTGGAATGCTTGGCAGCCGGTGGAGACATCTCTCTCCCAACGCTAAAT 528
DB 1510 GAGGGCGTCTGAGTCCCAAGTGTGAGACATGAGGCGCTCTTAATCTCTGGAT 1569
QY 529 TTCAACGAGCGCATCTTAAGAGAGACGCTGTGCGCCCGCAGAGATCAACATTTCTG 588
DB 1570 CTTAATGACACTCATCTCTGGAGATCAAGCTCAGAGCCCAATATCAATCTTTCTG 1629
QY 589 GATCAGAGTGAAGGGTCCCTGCTGTCGATGATCTTTGGATACCCCTGATGACCTGGAT 648
DB 1630 GACCAAGTGAAGGATCTATTCTCTGATGATACTTGGACAGTCCAGATGAATGAC 1689
QY 649 ATTAAGTGTGATGACATCGAGACCCCGGATGAGACGATCTCGTGGAGTCTCTGGGAAAT 708
DB 1690 ATCAATGTGTGATGAATCTTATATACCCCGGATGAGCAGATCTTTTGGATACACTGG 1745
QY 709 GGCAACGAATGAGTGGGAGACACACCCCGTGGCCACCGCCAGAACATGCGCGG 768
DB 1746 -----CCATGATCCACAGCCCAAGAAATCTTGCC 1776
QY 769 GACAGCGCGATCTATTGCGGACGCGACGAGGAGCGGCGCGCCCAACGCGCGC 828

DB 1777 CAAGAGTCAGAGTCTATTTCAGAAATATACGCGCGAAGAGAAACGCGAGACAAC ---CGG 1833
QY 829 CTGTGCGGACAGTATCATTCGGGAGCAGAGCAGCGGTATAGACTGACATGATCCCG 888
DB 1834 CTTTGAGGACAGTGGTCAATGGAACAAGAGCAGCGCATTTGACATGAAGGTCAATCGAG 1893
QY 889 CTTTACATGAAGTGGTCAACCCAGGAGGTACTACGGGAAAGGCTCAACGCCATCATC 948
DB 1894 CCTACAGAGAGTCAATTTCTCAGGAGGATATATGGGAGCGGTCTAATGCCATCAT 1953
QY 949 GTCTTGGAGCTGCTTCTTCCAGACAGACCTCCCGACTACCACTACATCATCATGAG 1008
DB 1954 GTGTTTGGCGCTGTTTCTGCGCAGACAGCAGTCCGGGCGATTACCACTATGTCATGGA 2013
QY 1009 AACCTCTTCTCTGATCATCAGCAGCTTAGAGCTCTGCTGGCTGAGAGCTACATGATC 1068
DB 2014 AATCTTTTCTATATGTAATAAGTACTTTTAGAGTTGNTGGTAGCTGAGAGCTATATGAT 2073
QY 1069 GTGTACCTGAACGCTGCCACGCCC -CGGCGGAGGATGCTCGAATCGGCTGCTGAAGAA 1127
DB 2074 GTGTACTTGAATGGTGCACACCCCAAGAGAGGAGATGCCAGGCTAGGCTGGATGAAGAA 2133
QY 1128 GTGTACCATGATGATCGACCGGAGGTTCGGGAAACCTGAAAGTCTTGTGATCATCGTCCA 1187
DB 2134 ATGCTACCATGATGATTCAGACAGCGTTGAGGAAGAAATTTGAAATCATTCATCATTTGTTCA 2193
QY 1188 CCCTCTGTGTTTCAATCGGACTGTGCTGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTT 1247
DB 2194 TCCATCTTGGTTTTCATCAGAACAAATCTTGTCTGTGACAGACCTTTTATAAGTTTCAAAAT 2253
QY 1248 CATCAACAAGATCCAGTACGTCAGACAGCTTGGAGACCTTGGAGCAACTCATCCCTATGGA 1307
DB 2254 CAGCAGTAAATTAATATGCTCAATAGCTTATCAGAACTCAGTGGGCTGATCCCAATGGA 2313
QY 1308 ACAGTCCAGATCCAGAGTGGCTCTGCAATACGAAGAGGAA 1350
DB 2314 TTGCATCCCAATTTCCAGAGAGCATCATCAAACTGGATGAAGAA 2356

RESULT 15

US-09-968-007A-729

; Sequence 729, Application US/09968007A
; Publication No. US20040115625A1

GENERAL INFORMATION:

; APPLICANT: Ebner, Reinhard

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signa

; FILE REFERENCE: 689290-71

; CURRENT APPLICATION NUMBER: US/09/968, 007A

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 172

; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 173

; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 278

; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 294

; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 295

; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: US/60/237, 316

; NUMBER OF SEQ ID NOS: 1001

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO: 729

; LENGTH: 5654

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-968-007A-729

Query Match 10.9%; Score 318.2; DB 3; Length 5654;
Best Local Similarity 63.2%; Pred. No. 8.7e-81;

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 03:46:55 ; Search time 810 Seconds
(without alignments)

7898.360 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gccggcctctgccgcct.....tctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA_New.*

- 1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2598.2	89.0	2759	9	US-11-072-512-997
2	1325.8	45.4	2625	9	US-11-072-512-986
3	243.6	8.3	623	6	US-09-925-065A-515679
4	241.4	8.3	403278	8	US-10-995-561-13421
5	241.4	8.3	560	6	US-09-925-065A-594866
6	240.8	8.3	138827	7	US-10-330-773-159
7	239.8	8.2	560	6	US-09-925-065A-594865
8	239.6	8.2	628	6	US-09-925-065A-55092
9	239.2	8.2	611	6	US-09-925-065A-482665
10	239.2	8.2	611	6	US-09-925-065A-563212
11	239.2	8.2	67126	8	US-10-995-561-13342
12	239.2	8.2	127340	12	US-11-112-908-35
13	239	8.2	483	6	US-09-925-065A-621024
14	239	8.2	1082	6	US-09-925-065A-84578
15	238.6	8.2	2649	6	US-09-925-065A-723891
16	238.4	8.2	547	6	US-09-925-065A-657480
17	238.4	8.2	623	6	US-09-925-065A-687639
18	238.4	8.2	623	6	US-09-925-065A-687640
19	238.2	8.2	624	6	US-09-925-065A-849333
20	238	8.2	180531	9	US-11-114-798-57

c 21	237.8	8.1	694	6	US-09-925-065A-75055	Sequence 75055, A
c 22	237.8	8.1	1288	6	US-09-925-065A-692997	Sequence 692997, A
c 23	237.6	8.1	611	6	US-09-925-065A-563211	Sequence 563211, A
c 24	237.4	8.1	482	6	US-09-925-065A-624538	Sequence 624538, A
c 25	237.4	8.1	482	6	US-09-925-065A-624539	Sequence 624539, A
c 26	237.4	8.1	570	6	US-09-925-065A-197384	Sequence 197384, A
c 27	237.2	8.1	621	6	US-09-925-065A-337721	Sequence 337721, A
c 28	236.6	8.1	883	6	US-09-925-065A-8768	Sequence 8768, A
c 29	236.2	8.1	682	6	US-09-925-065A-486749	Sequence 486749, A
c 30	236.2	8.1	883	6	US-09-925-065A-8769	Sequence 8769, A
c 31	236.2	8.1	187745	12	US-11-121-086-83	Sequence 83, Appl
c 32	236	8.1	600	6	US-09-925-065A-873038	Sequence 873038, A
c 33	236	8.1	661	6	US-09-925-065A-94635	Sequence 94635, A
c 34	236	8.1	1980	6	US-09-925-065A-92083	Sequence 92083, A
c 35	236	8.1	1980	6	US-09-925-065A-96162	Sequence 96162, A
c 36	236	8.1	171486	12	US-11-121-086-105	Sequence 105, App
c 37	236	8.1	179777	12	US-11-121-086-106	Sequence 106, App
c 38	235.8	8.1	199321	12	US-11-121-086-10	Sequence 10, Appl
c 39	235.6	8.1	661	6	US-09-925-065A-94636	Sequence 94636, A
c 40	235.6	8.1	3125	9	US-11-072-512-1355	Sequence 1355, A
c 41	235.6	8.1	40000	8	US-10-995-561-13512	Sequence 13512, A
c 42	235.4	8.1	196200	12	US-11-121-086-9	Sequence 9, Appl
c 43	235.2	8.1	574	6	US-09-925-065A-443115	Sequence 443115, A
c 44	235	8.1	553	6	US-09-925-065A-930747	Sequence 930747, A
c 45	235	8.1	580	6	US-09-925-065A-282862	Sequence 282862, A

ALIGNMENTS

RESULT 1
US-11-072-512-997
; Sequence 997, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTOKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 997
; LENGTH: 2759
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-997

Query Match 89.0%; Score 2598.2; DB 9; Length 2759;
Best Local Similarity 95.2%; Pred. No. 7.5e-25;
Matches 2756; Conservative 0; Mismatches 3; Indels 136; Gaps 2;

Db 2025 AATGACACCCAGTAACTAGAACTTCTCAAGCCCTTTAACTCAGATGTCAAGCCACC 2084
QY 2228 GGGCAAAACCCCGTCAATACCTCCCAAGGAATGAGATATGGAACCTCACTGTCTCCCC 2287
Db 2085 GGGCAAAACCCCGTCAATACCTCCCAAGGAATGAGATATGGAACCTCACTGTCTCCCC 2144
QY 2288 CAACCCAGGTCAGGCTGGGACACGCAAGCTGTTCGGGTTGGAAAGAGAGGCTCA 2347
Db 2145 CAACCCAGGTCAGGCTGGGACATGCAAGCTGTTCGGGTTGGAAAGAGAGGCTCA 2204
QY 2348 GAACTGGCTCTGAATAGCAGACCTAGCAGGAAGATACAGGATATCGGGGTTTG 2407
Db 2205 GAACTGGCTCTGAATAGCAGACCTAGCAGGAAGATACAGGATATCGGGGTTTG 2264
QY 2408 AGTGTTCAGAACTCAATTCGGGAAGATAAATCCAGTGCCTGGCCAGACCACTGCAAT 2467
Db 2265 AGTGTTCAGAACTCAATTCGGGAAGATAAATCCAGTGCCTGGCCAGACCACTGCAAT 2324
QY 2468 CAAGCTTGGACAGCGGTTCTGTTCGGGAGGCAAAATTCCTAGGAAAAGAGACA 2527
Db 2325 CAAGCTTGGACAGCGGTTCTGTTCGGGAGGCAAAATTCCTAGGAAAAGAGACA 2384
QY 2528 GACTTTTCTAATGTGTCTCAAAATGCGGATCACTGTGTCAGATGGAATCTAGAGCACTGAG 2587
Db 2385 GACTTTTCTAATGTGTCTCAAAATGCGGATCACTGTGTCAGATGGAATCTAGAGCACTGAG 2444
QY 2588 CTCCTGTCTCTGGAAGTATTTAAGAAAAGGCTGGGCCAGGCAAGATGCTCAAGCTGT 2647
Db 2445 CTCCTGTCTCTGGAAGTATTTAAGAAAAGGCTGGGCCAGGCAAGATGCTCAAGCTGT 2504
QY 2648 AATCCAGACTTTGGAGGCGGAGGCAAGCGGATCACTGAGGTGAGGAGTTTGAGAA 2707
Db 2505 AATCCAGACTTTGGAGGCGGAGGCAAGCGGATCACTGAGGTGAGGAGTTTGAGAA 2564
QY 2708 GCTGCGCAACATGGTGAACCTCATCTACTATAAATAAATAAATTAAGCAGGCTGG 2767
Db 2565 GCTGCGCAACATGGTGAACCTCATCTACTATAAATAAATAAATTAAGCAGGCTGG 2624
QY 2768 TGGCAGGTGCTGTATATCCAGCTACTTGGGAGGCTGAGGCATGAGATCACTTAAACCT 2827
Db 2625 TGGCAGGTGCTGTATATCCAGCTACTTGGGAGGCTGAGGCATGAGATCACTTAAACCT 2684
QY 2828 GAGAGCAGAGTTTACAGTGAGCCAGATCGTGCCAATGCAATTCAGGCTGGGCGACAGA 2887
Db 2685 GAGAGCAGAGTTTACAGTGAGCCAGATCGTGCCAATGCAATTCAGGCTGGGCGACAGA 2744
QY 2888 GCAAGACTCTGTCTC 2902
Db 2745 GCAAGACTCTGTCTC 2759

RESULT 2

US-11-072-512-986
; Sequence 986, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YUKI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYUKI
; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 986
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-986

Query Match 45.4%; Score 1325.8; DB 9; Length 2625;

Best Local Similarity 99.8%; Pred. No. 5.1e-10;

Matches 1338; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1562 CCCAGATCTCATCTGCTCATCTGATGCCAATCTTCCAGGGTGCAGCCCTCCG 1621
Db 1286 CTCAGATCTCATCTGCTCATCTGATGCCAATCTTCCAGGGTGCAGCCCTCCG 1345
QY 1622 TTCATCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAACATTTATTTTTTTT 1681
Db 1346 TTCATCTCTGAAACCCAGCATCTTTTTCAGCTGCTTGAACATTTATTTTTTTT 1404
QY 1682 AACGATGAGTATTTTGTGCTTCCAGAAAAGGGCCAGCTCTGAGCCCTCAACCTTCCA 1741
Db 1405 AACGATGAGTATTTTGTGCTTCCAGAAAAGGGCCAGCTCTGAGCCCTCAACCTTCCA 1464
QY 1742 CACTACGAACTCTCAGCCGAGGAAGGCAAGAGCGCAGGGGTGGCCGGTGGCGTGG 1801
Db 1465 CACTACGAACTCTCAGCCGAGGAAGGCAAGAGCGCAGGGGTGGCCGGTGGCGTGG 1524
QY 1802 GTGGCTCCGCTCTGCTCGCAGCCCTGCTGAGCTGAGCTGAGTACAGATTCAGAGCC 1861
Db 1525 GTGGCTCCGCTCTGCTCGCAGCCCTGCTGAGCTGAGCTGAGTACAGATTCAGAGCC 1584
QY 1862 TTCTCTGCTTGTGTCACCCGCTCCAGGTTGAGGCCACAGACACCCACCCCGGCTG 1921
Db 1585 TTCTCTGCTTGTGTCACCCGCTCCAGGTTGAGGCCACAGACACCCACCCCGGCTG 1644
QY 1922 GGTCTGCTCTCTTCTGCTGCTTTCCTCCAGAAATGCGGCTCAGACTAGAACTCAA 1981
Db 1645 GGTCTGCTCTCTTCTGCTGCTTTCCTCCAGAAATGCGGCTCAGACTAGAACTCAA 1704
QY 1982 CCCCCCTATGAGGGCCAGCTCTGGGGTAGCTCTGAGCTCGGACCTTATGTCCAAATTT 2041
Db 1705 CCCCCCTATGAGGGCCAGCTCTGGGGTAGCTCTGAGCTCGGACCTTATGTCCAAATTT 1764
QY 2042 CACACCCATGGTTTTCATTTGACCCGCTTCTGCTCATATATGACACCCAGCTCT 2101
Db 1765 CACACCCATGGTTTTCATTTGACCCGCTTCTGCTCATATATGACACCCAGCTCT 1824
QY 2102 TTGAGAGGATCAGAGCCCATTTGCAAGAAGAGCGGCTGCCAACCATCTTGTCTCCGA 2161
Db 1825 TTGAGAGGATCAGAGCCCATTTGCAAGAAGAGCGGCTGCCAACCATCTTGTCTCCGA 1884
QY 2162 TTGCAAAATGACCCCAAGTAACTTAGAACAATTTCAAGCCCTTTAACTCAGATGTCAA 2221
Db 1885 TTGCAAAATGACCCCAAGTAACTTAGAACAATTTCAAGCCCTTTAACTCAGATGTCAA 1944
QY 2222 GCCACCGGGCAAAACCCCGTCAATACCTCCCAAGGAATGAGATATGTGGACCTCACTG 2281
Db 1945 GCCACCGGGCAAAACCCCGTCAATACCTCCCAAGGAATGAGATATGTGGACCTCACTG 2004
QY 2282 CTCGCCCAACCCAGGCTCAGGCTGGGACACCGCCAAACGCTGTTCCGGGTGGAAACAGCAGA 2341
Db 2005 CTCGCCCAACCCAGGCTCAGGCTGGGACACCGCCAAACGCTGTTCCGGGTGGAAACAGCAGA 2064

RESULT 5

US-09-925-065A-594866
; Sequence 594866, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594866
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-594866

Query Match 8.3%; Score 241.4; DB 6; Length 560;
Best Local Similarity 85.1%; Pred. No. 7.6e+03;
Matches 280; Conservative 1; Mismatches 47; Indels 1; Gaps 1;

QY 2590 CCTGTCTCTGGAAGTATTAAAGAAAGGCTGGCCAGGACGATGGCTCAGCCTGTAA 2649
DB 1 CCAGGTGATGCAAGGTTCAAGAGTGTTCAGCCAGGCGATGGTGGCTCACACCTGTAA 60
QY 2650 TCCAG-ACCTTGGAGGCGGAGGCGGATCCTCAGGTGAGGAGTTTGAGACAG 2708
DB 61 TCCAGACACTTGGAGGCGGAGGCTGGTGGATCCTCAGGTGAGGAGTTCAAGACAC 120
QY 2709 CTGGCCAACTGGTGAACCTCATCTCTACTAAATAACAAATAGCCAGGCTGT 2768
DB 121 CTGGCCAACTGGAGAACCCCTCTCTACTAAATAACAAATAGCCAGGCTGT 180
QY 2769 GGCAGTGGCTGTAAATCCAGCTACTTGGAGGCTGAGCATGAGTAATCACTTAAACCTG 2828
DB 181 GGCAGTGGCTGTAAATCCAGCTACTTGGAGGCTGAGCATGAGTAATGCTTAAACCCA 240
QY 2829 AGAGGCAGAGGTTACAGTCAGGCAAGATCGTCCAGTCATTCAGCTGGGCGACAG 2888
DB 241 GAGGTGGAGTGTTCAGTGAGCCAGATCGTGCCTCAGCTCCAGCTGGGCGACAG 300
QY 2889 CAAGACTCTGTCTCAAAAAAATAAAAAA 2917
DB 301 CGAGATCTGTCTCAAAAAAATAAAAAA 329

RESULT 6

US-10-330-773-159/c
; Sequence 159, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:

; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 159
; LENGTH: 138627
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(138627)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-159

Query Match 8.3%; Score 240.8; DB 7; Length 138627;
Best Local Similarity 89.1%; Pred. No. 83;
Matches 271; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 2616 AGGCTGGCCAGGACGATGGCTCAGGCTGTAATCCAG-ACCTTGGAGGCCGAGGCA 2674
DB 108106 AGTGTGGCCAGGCGTGTGGCTCAGGCTGTAATCCAGCACTTTGGAGGCCGAGGCG 108047
QY 2675 GCGGATCACCTGAGGTGAGGAGTTTGAGAAACAGCCTGGCCAAACATGGTGAACCTCATC 2734
DB 108046 GGTGATCATCTGCACTGAGGAGTTGAGACAGCCTGCCCAAGATGGTGAACCCCGTC 107987
QY 2735 TCTACTAAAAATACAAAAATTAGCCAGGCTGGTGGAGGTCCTGTATCCAGCTACT 2794
DB 107986 TCTACTAAAAATTCAAAAATTAGCCAGGCTGGTGGAGGTCCTGTATCCAGCTACT 107927
QY 2795 TGGGAGGCTGAGGATGAGATCACTTAAACCTGAGGCGAGGCTTACAGTGAGCCAG 2854
DB 107926 CAGGAGGCTGAGGCGAGGAGATCGTTGAACCTGGAGCGGAGGTTCAGTGAGCTAAG 107867
QY 2855 ATCGTGCCACTGCATTCCAGCCTGGCCGACAGAGCAAGACTCTGTCTCAAAAAAATA 2914
DB 107866 GTCATGCCACTGCATCCAGCTGGCCGATAGAGCAAGCTGTCTCAAAAAAATA 107807
QY 2915 AAAA 2918
DB 107806 AAAA 107803

RESULT 7

US-09-925-065A-594865
; Sequence 594865, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594865
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-594865

Query Match 8.2%; Score 239.8; DB 6; Length 560;
Best Local Similarity 84.8%; Pred. No. 8e+03;
Matches 279; Conservative 1; Mismatches 48; Indels 1; Gaps 1;

QY 2590 CCTGTCTCTGGAAGTATTAAAGAAAGGCTGGCCAGGACGATGGCTCAGCCTGTAA 2649

Db 1 CCCAGTGCATGCAAGGTTCAAGAGTGGTTCAGCCAGCAGCATGGTGGCTCACACCTGTAA 60
QY 2650 TCCAG-ACCTTTGGGAGCGGAGCGAGGATCACCTGAGGTGAGGATTTGAGAACAG 2708
Db 61 TCCAGCACCTTTGGGAGCGGAGGTGGTGGATCACCTGAGGTGAGGATTTGAGAACAC 120
QY 2709 CTGCGCAACATGGTGAAACCTCATCTCTACTAAAAATACAAAAATAGCCAGCGTGGT 2768
Db 121 CTGCGCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAATAGCCAGCGTGGT 180
QY 2769 GCGAGTGCCTGTAATCCAGCTTACTTGGAGGCTGAGCATGAGATCACTTAAACCTG 2828
Db 181 GCGAGTGCCTGTAATCCAGCTTACTTGGAGGCTGAGCATGAGATCACTTAAACCTG 240
QY 2829 AGAGGAGAGGTACAGTACAGCAAGATCGTGCCTGCACTCCAGCCTGGGCGACAGAG 2888
Db 241 GAGGTGGATGTTTCARTAGCCAGATCGTGCCTGCACTCCAGCCTGGGCGACAGAG 300
QY 2889 CAAGACTCTGTCTCAAAAAA 2917
Db 301 CGAGATCTGTCTCAAAAAA 329

RESULT 8

US-09-925-065A-55092
; Sequence 55092, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55092
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-55092

Query Match 8.2%; Score 239.6; DB 6; Length 628;
Best Local Similarity 87.3%; Pred. No. 7.3e+03;
Matches 274; Conservative 0; Mismatches 39; Indels 1; Gaps 1;
QY 2605 TATTTAAGAAAGGCTGGGCGAGCAGATGGCTCAGCGCTGTAATCCAG-ACCTTTGGG 2663
Db 67 TAAAAA 126
QY 2664 AGCCGAGGCGCGGATCACCTGAGGTGAGGATTTGAGAAACAGCTGGCCCAACATGTT 2723
Db 127 AGCCGAGGCGAGCAGATCACTTTGAAGTCAGGAGTTTCGAGACCAGCCTGGCCCAACATGTT 186
QY 2724 GAAACCTCATCTCTACTAAAAATACAAAAATAGCCAGGCTGGTGGCAGGCTGCTGTAA 2783
Db 187 GAAACCTCATCTCTACTAAAAATACAAAAATAGCCGAGGCTGGTGGTGGTGGTGGTGGTGGT 246
QY 2784 TCCAGCTACTTGGGAGGCTGAGGCTAGGATGAGATCACTTAAACCTGAGAGGCGAGGTTAC 2843
Db 247 TCCAGCTACTTGGGAGGCTGAGGCGAGGAGATCGCTTGAACCTAGGAGGCGAGGTTGC 306

QY 2844 ACTGAGCCAGATCGTGCCACTGCACTCCAGCTTGGGCGACAGAGCAAGACTCTGTCTCA 2903
Db 307 ACTGAGCGAGATCGTGCCACTGCACTCCAGCTTGGGCGACAGAGTGTGTCTCA 366
QY 2904 AAAAAA 2917
Db 367 AAAAAA 380

RESULT 9

US-09-925-065A-482665
; Sequence 482665, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482665
; LENGTH: 562
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-482665

Query Match 8.2%; Score 239.2; DB 6; Length 562;
Best Local Similarity 87.5%; Pred. No. 8.1e+03;
Matches 273; Conservative 0; Mismatches 38; Indels 1; Gaps 1;
QY 2608 TTAAGAAAGGCTGGGCGAGCAGATGGCTCAGCGCTGTAATCCAG-ACCTTTGGGAGG 2666
Db 159 TAAGGTATAATCTAGGCCAGGACCGTGTCTCATCTGTAATCCAGAACTTTGGGGG 218
QY 2667 CCGAGGAGGCGGATCACCTGAGGTGAGGATTTGAGAACAGCCTGGCCCAACATGGTAA 2726
Db 219 CCGAGGCTGGTGGATCACCTGAGGTGAAGAGATTGAGACCATCTCTGGCCCAACATGGTAA 278
QY 2727 ACCTCATCTCTACTAAAAATACAAAAATAGCCAGGCTGGTGGCAGGCTGCTGTAATCC 2786
Db 279 ACCCTGTCTCTACTAAAAATACAAAAATAGCCAGGCGGGTGGCAGGACCTGTAATCC 338
QY 2787 CAGCTACTTGGGAGGCTGAGGATGAGATCACTTAAACCTGAGAGGCGAGGTTACAGT 2846
Db 339 CAGCTACTTGGGAAGCTGAGGCAAGAGATCACTTGAACCCGGGAGGCGAGGTTGAGT 398
QY 2847 GAGCAAGATCGTGCCACTGCAATTCAGCCTGGGCGAGCAGAGCAAGACTCTGTCTCAAAA 2906
Db 399 GATCCAAAGTATGTCACCTGCACTCCAGCTGGGTGAGCAGAGTGAGACTCCGTCTCAAAA 458
QY 2907 AAAAAA 2918
Db 459 AAAAAA 470

RESULT 10

US-09-925-065A-563212/c
; Sequence 563212, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

Db 124909 GAGCCAGATCATGCCACTGCAGCTCCAGCTGGGCAACAGAGTGAGACTCCATCTTAAAA 124850

QY 2907 AAAAAAAAAAAAA 2918

Db 124849 AAAAAAAAAAAAA 124838

RESULT 13

US-09-925-065A-621024/c

Sequence 621024, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 621024

LENGTH: 483

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-621024

Query Match 8.2%; Score 239; DB 6; Length 483;
Best Local Similarity 89.0%; Pred. No. 9.2e+03;
Matches 268; Conservative 1; Mismatches 31; Indels 1; Gaps 1;

QY 2619 CTGGGCGCAGCATGGCTCAGCCTGTAAATCCAG-ACCTTGGGAGCGGCGAGCG 2677

Db 483 CTGGGCGAGTGCGAGTGGCTCATGCTGTAAATCCAGCACTTGGGAGGCTGAGGCGGC 424

QY 2678 GATACCTCAGGTGAGGAGTTTGAGAACAGCCTGGCCAAACATGGTGAACCTCATCTCT 2737

Db 423 TGATCACTGAGTCCAGGAGTTCGAGCCAGCTGGCCAAACATGGTGAACCTCATCTCT 364

QY 2738 ACTAAAAATACAAAAATTAGCCAGGCGTGGTGCAGGCTGTGTAAATCCAGCTACTTGG 2797

Db 363 ACTAAAAATACAAAAATTAGTCTGGTGTGGTGCAGGCTGTGTAAATCCAGCTACTTGG 304

QY 2798 GAGGCTGAGCATGAGATCACTTAACTTGAGGCGAGGTTACAGTGAGCCAGATC 2857

Db 303 GAGGCTGAGCATGAGATCACTTAACTTGAGGCGAGGTTACAGTGAGCTGAGATG 244

QY 2858 GTGCCACTGATTCAGCCTGGGCGACAGAGCAAGACTCTGTCTCAAAAAA 2917

Db 243 GTGCCACTGATTCAGCCTGGGCGACAGAGCAAGACTCTGTCTCAAAAAA 184

QY 2918 A 2918

Db 183 A 183

RESULT 14

US-09-925-065A-84578/c

Sequence 84578, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 84578

LENGTH: 1082

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-84578

Query Match 8.2%; Score 239; DB 6; Length 1082;

Best Local Similarity 87.5%; Pred. No. 4.7e+03;

Matches 260; Conservative 1; Mismatches 36; Indels 0; Gaps 0;

QY 2622 GGCAGGCGACGATGGCTCAGCCTGTAAATCCAGACTTTGGGAGGCCGAGCGCGAT 2681

Db 412 GGCAGGCGCGGTGGCTCAGTCTGTAAATCCAGACTTTGGGAGGCCGAGCGCGAT 353

QY 2682 CACCTGAGGTGAGGAGTTTGAGAACAGCCTGGCCAAACATGGTGAACCTCATCTCTACTA 2741

Db 352 CGCCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAGCATGGTGAACCTCATCTCTACTA 293

QY 2742 AAAATACAAAATTTAGCCAGGCGTGGTGCAGGCTGCTGTAAATCCAGCTACTTGGGAGG 2801

Db 292 AAAATACAAAATTTAGCCAGGCGATGGTGCAGGCTGCTGTAAATCCAGCTACTTGGGAGG 233

QY 2802 CTGAGGCGATGAGAACTCACTTAAACCTGAGAGCGAGGTTACAGTGAGCAAGATCGTGC 2861

Db 232 CTGAGGCGATGAGAACTCACTTAAACCTGAGAGCGAGGTTACAGTGAGCAAGATCGTGC 173

QY 2862 CACTGCATTCAGCCTGGGCGACAGCAAGACTCTGTCTCAAAAAA 2918

Db 172 CATTGCACTCCAGCCTGGGCGAAAGAGCGAGACTCTGTCTCAAAAAA 116

RESULT 15

US-09-925-065A-723891/c

Sequence 723891, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 723891

LENGTH: 2649

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-723891

Query Match 8.2%; Score 238.6; DB 6; Length 2649;
Best Local Similarity 88.5%; Pred. No. 2.3e+03;
Matches 270; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY	2615	AAGCTGGCCAGGCGATGGCTCAGCCCTGTAATCCAG-ACCTTGGAGGCCGAGGC	2673
Db	2260	AAAGATGGGCCAGGCGTGGCTCAGCCCTGTAATCCAGCACTTTGGAGGCCGAGGC	2201
QY	2674	AGCGGATCACCTGAGGTGAGGAGTTGAGAACAGCCTGGCCAAACATGGTGAACCTCAT	2733
Db	2200	GGGTGGATCACTTGAAGTCAAGGAGTTGAGACCAGCCTGGCCAAACATGGTGAATCCCGT	2141
QY	2734	CTCTACTAAATAACAAAAATTAGCCAGGCGTGGTGCGAGGTGCCTGTATATCCAGCTAC	2793
Db	2140	CTCTACTAAAGATACAAAAATTAGCCAGGCTGGTGCTGGGCGCTGTATATCCAGCTAC	2081
QY	2794	TTGGGAGGCTGAGGCGATGAGAAATCACTTAACTGAGAGGCGAGAGGTTACAGTGAGCCAA	2853
Db	2080	TCGGGAGACTGAGGCGAGGAGAACTCACTTGAACCTGGAGGCGAGGTTGCAGTGAGCTGA	2021
QY	2854	GATCGTGCCACTGCATTCAGCCTGGCGGACAGAGCAAGACTCTGTCTCAAAAAA	2913
Db	2020	GATTGTGACACTGCATTCAGCCTGGCGGACAAACAGACTCCATCTCAAAAAA	1961
QY	2914	AAAAA 2918	
Db	1960	AAAAA 1956	

Search completed: March 5, 2006, 07:06:18
Job time : 815 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2006, 03:31:06 ; Search time 11150 Seconds
(without alignments)

12244.368 Million cell updates/sec

Title: US-10-699-941-3

Perfect score: 2918

Sequence: 1 gcggcgctctgccgcctcct.....tctcaaaaaaaaaaaaaa 2918

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gsa1:*

10: gb_gsa2:*

11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2782.6	95.4	3253	4	CR857898
2	940	32.2	1362	11	DQ053107
3	857.8	29.4	2104	4	AK038483
4	857.8	29.4	2601	4	AK028298
5	857.8	29.4	2603	4	AK039411
6	857.8	29.4	3679	4	AK038990
7	831	28.5	2148	4	AK086723
8	783	26.8	996	11	DQ053108
9	748	25.6	748	7	CN420201
10	730	25.0	865	2	BG912726
11	696.4	23.9	699	7	CV024962
12	693.4	23.8	714	8	DR002011
13	653	22.4	876	2	BG913524
14	648.4	22.2	650	1	AL138250
15	639	21.9	639	7	CN420203
16	635.4	21.8	884	5	BUS21943
17	632.4	21.7	688	7	CR790459
18	618	21.2	630	6	CB153985
19	595.6	20.4	842	2	BF345891
20	585	20.0	609	2	BG910297
21	578.6	19.8	800	6	CA315557
22	575.4	19.7	609	2	BG819886

23	570.6	19.6	598	7	CR536713	CR536713	DKFZp459E
24	565.6	19.4	606	1	AW961890	AW961890	EST373963
25	544.8	18.7	1022	2	BF530295	BF530295	602071474
26	544	18.6	1043	2	BG913501	BG913501	602811232
27	539	18.5	577	8	DR002579	DR002579	TC107021
28	538.2	18.4	689	3	BQ179996	BQ179996	UI-M-EW0-
29	525.8	18.0	946	2	BF527108	BF527108	602039553
30	520.6	17.8	548	7	CR773543	CR773543	DKFZp459E
31	502.6	17.2	673	7	CN219420	CN219420	WLA005H09
32	496.8	17.0	729	6	CA750371	CA750371	UI-M-FY0-
33	483.4	16.6	509	7	CR628631	CR628631	DKFZp459L
34	482.6	16.5	494	3	BQ312270	BQ312270	QV1-EN039
35	481.8	16.5	637	3	BQ044387	BQ044387	UI-M-EH0-
36	459	15.7	608	2	BF527542	BF527542	602040445
37	441.4	15.1	1471	8	DN717620	DN717620	CNB119-F0
38	440.6	15.1	495	1	AI498192	AI498192	tm90912.x
39	422.4	14.5	576	7	CN420200	CN420200	170005322
40	410.6	14.1	521	6	CB720956	CB720956	AMGNNUC:N
41	409.8	14.0	1087	8	DN689249	DN689249	CGX63-E08
42	405.8	13.9	1402	8	DN723748	DN723748	CNB150-CO
43	390.2	13.4	752	6	CB245546	CB245546	UI-M-FY0-
44	390.2	13.4	1169	8	DN735716	DN735716	CNB82-H02
45	386.4	13.2	461	1	AI480208	AI480208	tm72f02.x

ALIGNMENTS

RESULT 1	CR857898	CR857898	3253 bp	mRNA	linear	HTC 17-APR-2005
LOCUS	Pongo pygmaeus mRNA; cDNA DKFZp459H1854 (from clone DKFZp459H1854).					
DEFINITION	CR857898					
ACCESSION	CR857898.1	GI:55726776				
VERSION	HTC.					
KEYWORDS	Pongo pygmaeus (orangutan)					
SOURCE	Pongo pygmaeus					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.					
REFERENCE	1 (bases 1 to 3253)					
AUTHORS	Koehrer, K., Beyer, A., Mewes, H.W., Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.					
CONSRMT	The German cDNA Consortium					
TITLE	Direct Submission					
JOURNAL	Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY					
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BHPZ (Biomedical Research Center at the Heinrich-Heine-University, Duesseidorf/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp459H1854) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp459H1854 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/					
FEATURES	Location/Qualifiers					
source	1. 3253					
	/organism="Pongo pygmaeus"					
	/mol_type="mRNA"					
	/db_xref="taxon:9600"					
	/clone="DKFZp459H1854"					
	/tissue_type="cortex"					
	/clone_lib="459 (synonym: pcorl). Vector pSport1_Sfi; host DH10B; sites SfilA + SfilB"					
	/dev_stage="adult"					
	/note="hypothetical protein (Homo sapiens)"					
gene	1. 3253					
	/gene="DKFZp459H1854"					
CDS	408..1523					
	/gene="DKFZp459H1854"					

```
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH90150.1"
/db_xref="GI:5572677"
/db_xref="GOA:Q5RDK7"
/db_xref="InterPro:IPR001251"
/db_xref="UniProt/TrEMBL:Q5RDK7"
/translation="MGTTBATRMENVKSEWQDEDLPRPLPEETGVELLGSPVEDT
SSPNTLNFNGHKKRLTLVAPINISLQSGSLLSDPLDTDDLDINVDIDETPD
ETDSLEFLNGHLEWEDDTPVATKNMPSADLFGDGTEDGSAANGRLWRVTIG
EOHRDLHMKYKPVVTHGVYGEGLNAILVPAACFLPDSLDYHINENFLYV
ISLELLVADYIVYLNGATPRRMPGSLWAKCTQMDRLKRLNKLKSLIIVHPSWF
INTVLASIPFISVFKINKIQVHSLIEDLEQLIPMEHVQIPDVLQYBEERLARRES
TRPQEFVLPRSEKPEVAPVENRSALVSEDOEISMS"
```

ORIGIN

```
Query Match          95.4%; Score 2782.6; DB 4; Length 3253;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 2840; Conservative 0; Mismatches 74; Indels 3; Gaps 1;
```

QY	1	GCGGAGCCTCTGCGAGCCTCTGAGCTGGGAAGACGACGTACCTCGGAGGCGAGGCGCGCA	60
DB	36	GCGGAGCCTCTGCGAGCCTCTGAGCTGGGAAGACGACGTACCTCGGAGGCGAGGCGCGCA	95
QY	61	GCGGCGCGCGATGAGAGGGGGCGACGCCGACGCCCGCGCTGGGGAGCCCAACCGCTAAC	120
DB	96	GCGGCGCGCGATGAGAGGGGGCGACGCCGACGCCCGCGCTGGGGAGCCCAACCGCTAAC	155
QY	121	CTGTGACCCACCCACCCCTCGACAAAGAGCTGGGGGGCGCTGGCCAGTGCCTCGG	180
DB	156	CTGTGACCCACCCACCCCTCGACAAAGAGCTGGGGGGCGCTGGCCAGTGCCTCGG	215
QY	181	TGACCTTCTCGGATGACGAATCCGCCCTCGGAGCATCTCTTCTCTCTAGGCTCTGAA	240
DB	216	TGACCTTCTCGGATGACGAATCCGCCCTCGGAGCATCTCTTCTCTCTAGGCTCTGAA	275
QY	241	GCGCGGGGAGCGTGAGCGATGCCAGTGCACCCCGGGCAGGGCTCGCTTTGTTGCCA	300
DB	276	GCGCTGGGAGCGTGAGCGATGCCAGTGCACCCCGGGCAGGGCTCGCTTTGTTGCCA	335
QY	301	GTAAGGAGGAGGCTGTCTCAGCTGACAGGGGTATCCCTGCTTCAAGCCAGTGCCTC	360
DB	336	GTAAGGAGGAGGCTGTCTCAGCTGACAGGGGTATCCCTGCTTCAAGCCAGTGCCTC	395
QY	361	TTCCAGCTCCCATGGGACCAACCGACGCTCCGATGGMAAACGTGGACGTGAAG	420
DB	396	TTCCAGCTCCCATGGGACCAACGTGGACGCTCCGATGGMAAACGTGGACGTGAAG	455
QY	421	GAGGAATGGCAGGACGAAGATCTTCCAGGCCACTCCAGAGAGACGGGGTGGAACTG	480
DB	456	GAGGAATGGCAGGACGAAGATCTTCCAGGCCACTCCAGAGAGACGGGGTGGAAATG	515
QY	481	CTTGGCAGCCCGTGGAGACATCTCTCTCCCAACACGCTAAATTTCAACGGGCG	540
DB	516	CTTGGCAGCCCGTGGAGACATCTCTCTCCCAACACGCTAAATTTCAACGGGCG	575
QY	541	CATCGTAAGAGGAAGACGCTGTGGGCCCCAGAGATCAACATTTCTCGGATCAGAGTGAG	600
DB	576	CATCGTAAGAGGAAGACGCTGTGGGCCCCAGAGATCAACATTTCTCGGATCAGAGTGAG	635
QY	601	GGGTCCCTGCTGTCGATGACTTTCTTGGATATACCCCTGATGACTTGGATATTAACGTGGAT	660
DB	636	GGGTCCCTGCTGTCGATGACTTTCTTGGATATACCCCTGATGACTTGGATATTAACGTGGAT	695
QY	661	GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTCTCTGGGATGCGCAACGACTG	720
DB	696	GACATCGAGACCCCGATGAGACCGACTCGCTGGAGTCTCTGGGATGCGCAACGACTG	755
QY	721	GAGTGGGAAGACGACACCCCGTGGCCACCGCAAGAACATGCCGGGACGACGCGGAT	780
DB	756	GAGTGGGAAGATGACACCCCGTGGCCACTGCGCAAGAACATGCCGGGACGACGCGGAT	815
QY	781	CTATTGTTGGGACGGCAGGAGGACGGGACGGCGCCGCAACGGGGCGCTGTGGGCGGACA	840

DB	816	CTATTGTTGGGACGGCAGCAGGAGACGGCAGTGCAGCAACGGGCGCTGTGGCGGACG	875
QY	841	GTGATCATCGGGAGCAAGAGCACCGTATAGACCTGCAATGATCCGGCCTTACATGAAA	900
DB	876	GTGATCATCGGGAGCAAGAGCACCGTATGACCTGCAATGATCCGGCCTTACATGAAA	935
QY	901	TGTGTCAACCCACCGAGGAGTACTACGGCAAGGCGCTCAACGCCATCATCGTCTTCCAGCC	960
DB	936	TGTGTCAACCCACCGAGGAGTACTATGGCGAAGGCGCTCAATGCCATCATCGTCTTCCAGCC	995
QY	961	TGCTTCTTTCAGACAGCAGCGCTCCCGACATACCACTACATCATCATGAGAAACCTCTTCTG	1020
DB	996	TGCTTCTTTCAGACAGCAGCGCTCCCGACATACCACTACATCATCATGAGAAACCTCTTCTG	1055
QY	1021	TACGTCAATCAGCAGCTTAGAGCTCTGCTGGCTGAGGACTACATGATCGTGTACTCTGAAC	1080
DB	1056	TACGTCAATCAGCAGCTTAGAGCTCTGCTGGCTGAGGACTACATGATCGTGTACTCTGAAC	1115
QY	1081	GGTGCCACGCCCGCGGAGGATGCTCGAATCGGCTGGCTGAAGAGTGTCTACCAAGATG	1140
DB	1116	GGTGCCACGCCCGCGGAGGATGCTCGAATCGGCTGAAGAGTGTCTACCAAGATG	1175
QY	1141	ATCGACCGGAGGTTGCGGAAAAAAGCTGAAGTCTCTTGTATCATCGTCCACCCCTCGTGGTTC	1200
DB	1176	ATCGACCGGAGGTTGCGGAAAAAAGCTGAAGTCTCTTGTATCATCGTCCACCCCTCGTGGTTC	1235
QY	1201	ATTCCGAGCTGTGCTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCATCAACAGATC	1260
DB	1236	ATTCCGAGCTGTGCTGGCCATCTCTCGCCCTTTCATCAGCGTCAAGTTCATCAACAGATC	1295
QY	1261	CAGTACGTGTCACAGCTTGGAAAGACCTGGAGCAACTCATCTTATGAAACACGCTCCAGATC	1320
DB	1296	CAGTATGTGTCACAGCTTGGAAAGACCTGGAGCAACTCATCTTATGAAACACGCTCCAGATC	1355
QY	1321	CCAGACTCGCTCTGCAATACGAAGAGAAAGACTGAAGGCGCAGAGGAGGAGCGCAGG	1380
DB	1356	CCAGACTCGCTCTGCAATACGAAGAGAAAGGCTGGAGGCGCAGAGGAGGAGCGCAGG	1415
QY	1381	CCCCAGCCGGAGTTGTGCTGCCAGGCTCTGAGAGAGCCAGAGGTGGCACCAGTGAA	1440
DB	1416	CCCCAGCCGGAGTTGTGCTGCCAGGCTCTGAGAGAGCCAGAGGTGGCACCAGTGAA	1475
QY	1441	AACAGCTCTGCTCTGCTCTCAGAGATCAGGAAACAGCATGCTCTGAGGCGCAGCTGAGC	1500
DB	1476	AACAGCTCTGCTCTGCTCTCAGAGATCAGGAAATAGCATGCTCTGAGGCGCAGCTGAGC	1535
QY	1501	ATAACAAAGGACATGGAAGAAAGATTCAGATGCGCAAGAAACCTCTGTGACAGCCCACTG	1560
DB	1536	ATAACAAAGGACATGGAAGAAAGATTCAGATGCGCAAGAAACCTCTATCAGAGCGCCACTG	1595
QY	1561	GCCCCAGATCATCTCGCTCATCTCAGTCCCAATCTTCCAGGGTGGCAGCCCTCC	1620
DB	1596	GCCCCAGATCATCTCGCTCATCTCAGTCCCAATCTTCCAGGGTGGCAGCCCTCC	1655
QY	1621	GTTCACTCTCTGAAAACCCAGCATCTTTTTCAGCTGCTTGAAGAAACATTTGATTTTTTTT	1680
DB	1656	GTTCACTCTCTGAAAACCCAGCATCTTTTTCAGCTGCTTGAAGAAACATTTGATTTTTTTT	1712
QY	1681	TAAAGATGAGTATTTGTGCGTTTCAGAAAAGGGGCCAGCTCTGAGCCCTCACCCCTTCC	1740
DB	1713	TAAAGATGAGTATTTGTGCGTTTCAGAAAAGGGGCCAGCTCTGAGCCCTCACCCCTTCC	1772
QY	1741	ACACTCAGAACTCTCAGCCGAGGAGCAAGAGCGCAGGGGTGGCCCGGTCGCGTC	1800
DB	1773	ACACTCAGAACTCTCAGCCGAGGAGCAAGAGCGCAGGGGTGGCCCGGTCGCGTC	1832
QY	1801	GGTGGCTCTCCGCTCTCTGCTCGCAGCCCTGCTGGTTCAGAGCTGGATACAGATTTCAAGACC	1860
DB	1833	GGTGGCTCTCCGCTCTCTGCTCGCAGCCCTGCTGGTTCAGAGCTGGATACAGATTTCAAGACC	1892
QY	1861	CTTCTCTTGTGTCAACCCGCTCAGGTTGAGGCCACAGACCCACCGCCACCCCGGCT	1920

Db 421 TCCTGCTGCTCCGATGATCTTCTGGATACCCCTGATGACCTCGATATTAACGTGGATGAC 480
QY 664 ATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAAATGGCAACGAACTGGAG 723
Db 481 ATCGAGACCCCGATGAGACCGACTCGCTGGAGTTCTGGGGAAATGGCAACGAACTGGAG 540
QY 724 TGGGAGACACACCCCGTGGCCACCGCCCAAGAACATGCCCCGGGAGACGCGGATCTA 783
Db 541 TGGGAGACACACCCCGTGGCCACCGCCCAAGAACATGCCCCGGGAGACGCGGATCTA 600
QY 784 TTTGGGACGCGACGACGAGGAGCGCAGCGCCGCAACGCGGCGCTGTGGGCGACAGTG 843
Db 601 TTTGGGACGCGACGAGGAGCGCAGCGCCGCAACGCGGCGCTGTGGGCGACAGTG 660
QY 844 ATCATCGGGAGCAAGACGACCGTATAGACTCTGCATGATCCGGCTTACATGAAAGTG 903
Db 661 ATCATCGGGAGCAAGACGACCGTATAGACTCTGCATGATCCGGCTTACATGAAAGTG 720
QY 904 GTCAACCCAGGAGGTACTAGCGGAGGCTCAAGCCCATCATCGTCTTCGAGCGCTGC 963
Db 721 GTCAACCCAGGAGGTACTAGCGGAGGCTCAAGCCCATCATCGTCTTCGAGCGCTGC 780
QY 964 TTCCTTCCAGACGACGCTCCCGACTACCACTACATGATGAGAACCTCTTCCGTGAC 1023
Db 781 TTCCTTCCAGACGACGCTCCCGACTACCACTACATGATGAGAACCTCTTCCGTGAC 840
QY 1024 GTCATCAGCAGCTTAGAGCTCTCGTGGCTGAGGACTACATGATCGTGTACCTGAAACGGT 1083
Db 841 GTCATCAGCAGCTTAGAGCTCTCGTGGCTGAGGACTACATGATCGTGTACCTGAAACGGT 900
QY 1084 GCCACGCCCGCGGAGGATGCTCGAATCGGTGCTGAAGAGTGTACCAAGATGATC 1143
Db 901 GCCACGCCCGCGGAGGATGCTCGAATCGGTGCTGAAGAGTGTACCAAGATGATC 960
QY 1144 GACCGGAGTTGCGGAAACCTGAAGTCTTTCATCATGTCACCCCTCGTGTTCATT 1203
Db 961 GACCGGAGTTGCGGAAACCTGAAGTCTTTCATCATGTCACCCCTCGTGTTCATT 1020
QY 1204 CGGACTGTGCTGCCATCTCTCGCCCTTTCATCAGCGCTCAAGTTCATCAACAAGATCCAG 1263
Db 1021 CGGACTGTGCTGCCATCTCTCGCCCTTTCATCAGNNNNNNNNNNNNNNNNNNNNNN 1080
QY 1264 TAGCTGCACAGCTTGAAGACCTGGAGCACTCATCCCTATGGAACACGCTCCAGATCCA 1323
Db 1081 NNN 1140
QY 1324 GACTGGCTCTGCAATACGAGGAGGAACTGAAGCCAGGAGGAGGAGCGGAGGCC 1383
Db 1141 NNN 1200
QY 1384 CAGCCGAGTTTGTGCTGCCAGGTCTGAAGAGAGCCAGAGTGGCCACAGTGGAAAC 1443
Db 1201 CAGCCGAGTTTGTGCTGCCAGGTCTGAAGAGAGCCAGAGTGGCCACAGTGGAAAC 1260
QY 1444 AG 1445
Db 1261 AG 1262

RESULT 3
AK038483

LOCUS AK038483.1 GI:26332592
DEFINITION Mus musculus adult male hypothalamus cDNA, RIKEN full-length
enriched library, clone:A230019K19 product:weakly similar to
BCL2/ADENOVIRUS E1B 19-KDA PROTEIN-INTERACTING PROTEIN 2 [Mus
musculus], full insert sequence.

ACCESSION

AK038483

VERSION

AK038483.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

AUTHORS

TITLE

JOURNAL

PUBMED

Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

10349636

2 Carninci, P., Shibata, Y., Hayashizaki, Y., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kitsuura, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multiplexed sequencer

Genome Res. 10 (11), 1757-1771 (2000)

11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2104)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayashizaki, T., Hiramoto, K., Hiraoka, T., Hirozane, T.,

Hori, F., Imokawa, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, F.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,

Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,

URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

1. 2104

/organism="Mus musculus"

/mol_type="mRNA"

/strains="CS7BL/6J"

/db_xref="FANTOM_DB:A230019K19"

/taxon="10090"

/clone="A230019K19"

REFERENCE	1	Carninci, P. and Hayashizaki, Y.	
AUTHORS		High-efficiency full-length cDNA cloning	
TITLE		Meth. Enzymol. 303, 19-44 (1999)	
JOURNAL		10349636	
PUBLISHED			
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
AUTHORS		Normalisation and subtraction of cap-trapper-selected cDNAs to	
TITLE		prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	
PUBLISHED		11042159	
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Taghio, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, K., Izawa, M., Ohara, E., Watanabe, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
AUTHORS		RIKEN integrated sequence analysis (RISA) system--384-format	
TITLE		sequencing pipeline with 384 multicapillary sequencer	
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	
PUBLISHED		11076861	
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
AUTHORS		Functional annotation of a full-length mouse cDNA collection	
TITLE		Nature 409, 685-690 (2001)	
JOURNAL			
PUBLISHED			
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	
AUTHORS		Analysis of the mouse transcriptome based on functional annotation	
TITLE		of 60,770 full-length cDNAs	
JOURNAL		Nature 420, 563-573 (2002)	
PUBLISHED		6 (bases 1 to 2601)	
REFERENCE		Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tonari, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	
AUTHORS		Direct Submission	
TITLE		Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of	
JOURNAL		Physical and Chemical Research (RIKEN), Laboratory for Genome	
PUBLISHED		Exploration Research Group, RIKEN Genomic Sciences Center (GSC),	
REFERENCE		RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,	
AUTHORS		Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,	
TITLE		URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,	
JOURNAL		Fax: 81-45-503-9216)	
PUBLISHED		cDNA library was prepared and sequenced in Mouse Genome	
REFERENCE		Encyclopedia Project of Genome Exploration Research Group in Riken	
AUTHORS		Genomic Sciences Center and Genome Science Laboratory in RIKEN.	
TITLE		Division of Experimental Animal Research in Riken contributed to	
JOURNAL		prepare mouse tissues.	
PUBLISHED		Please visit our web site for further details.	
REFERENCE		URL: http://genome.gsc.riken.jp/	
AUTHORS		URL: http://fantom.gsc.riken.jp/	
TITLE		Location/Qualifiers	
JOURNAL		1. .2601	
PUBLISHED		/organism="Mus musculus"	
REFERENCE		/mol_type="mRNA"	
AUTHORS		/strain="C57BL/6J"	
TITLE		/db_xref="FANTOM DB:3322401A10"	
JOURNAL		/db_xref="taxon:10090"	
PUBLISHED		/clone="3322401A10"	
REFERENCE		/tissue_type="head"	
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBLISHED	</		

/clone_lib=RIKEN full-length enriched mouse cDNA library
/dev_stage="15 days embryo"
1. 2148
/notes="weakly similar to BCL2/ADENOVIRUS E1B 19-KDA
PROTEIN-INTERACTING PROTEIN 2 [Mus musculus]
(SWISSPROT|054940, evidence: PASTY, 52.5%ID, 94.7%length,
match=972)"

misc_feature

ORIGIN

Query Match 28.5%; Score 831; DB 4; Length 2148;
Best Local Similarity 74.9%; Pred. No. 4.7e-117;
Matches 1152; Conservative 0; Mismatches 340; Indels 47; Gaps 7;
64 GGGCGCGATGAGAGGGGGCGGCGGAGCCCGGCGCTGGGAGGCCACCGCTAACCCCT 123
DB |||||
89 GGGCTCGATGGCAGGGGAAACAGTGACTAAAGGGGACAGGGGTCTGCTACTAACCCA 148
QY 124 GCACCCACCCACCCCTGACAAAGAGCTGGGGGGCTGGCCACGTCGCCCTGGGTGA 183
DB |||||
149 CGGCGCGGCTTCTACAGTTGGGCGGACAGGTGTCTGTGGCCACGTCGCCCTGGGTGA 208
QY 184 CTTCTCTCGATGACAGATCCGCGCCCTGGGAGCATCTCTCTCTCTAGGCTCTGAAGGC 243
DB |||||
209 CTTCTCTCAGATGGA-----CTTGGCCCTGAGCAT 240
QY 244 CGGGGAGCGTGAGCGATGCCAGCTGCACCGGGCAGGGCTGCGCTTTTGTTCAGTA 303
DB |||||
241 CTTCTCACCGGCCCTCGCTGGGTACCATCGAGAAATGCCCGCTTTGTGTCCAAGTG 300
QY 304 AG-GAGGAGGCTGTCTCAGCTGCAGAG-----GGGTATCCCTGCTTCAAGCC 352
DB |||||
301 ACAGCCGAGGCGAGCTTCGCTTACCGAGATCATCTTCTGGGTCAACCCGAGTTTCAGAC 360
QY 353 AGTGCCCTTTCCAGCTCCCATGGGA-CCACCGAAGCCACGCTCCGGATGGAACGTCG 411
DB |||||
361 CACTTCTTTTCAGCTCTCATGGAAACCCACAGAAGCTACATAAGGATGGAATATGTC 420
QY 412 GACGTGAAGGAGGAATGGCAGGACGAGATCTTCCAGGCCCATCTCCAGAGACGCGGG 471
DB |||||
421 GACGTGAGGATGATGGCAGGATGAGGATCTGCCAGACCGCTCCAGAGACACCGGG 480
QY 472 GTGGAACTCTGGGAGCCCGGTGGGAAGACATCTCTCTCCCAACAGCTTAATTC 531
DB |||||
481 GTGGAGCGCTGGGTGGCGCAGTGGAAGACTCTCTCTCACTCCCTCCACCTGGAATTG 540
QY 532 AACGAGCGCATCGTAAGAGGAGACGCTGTGGGCCCGCCAGAGATCAACATTTCTGTGAT 591
DB |||||
541 AGCGGAGCATCGAAAGAGAACGCTGGTGGCTCCAGAGATCAACATCTCCCTGGAC 600
QY 592 CAGAGTGAGGGGTCCCTGTCTCCGATGACTTCTTGGATACCCCTGATGACCTGGATAT 651
DB |||||
601 CAAAGCGAGGCTCTCTGTCTCGACGACTTCTCGACACACCTGATGACCTGACATC 660
QY 652 AACGTGGATGACATCAGACCCCGATGAGACGACTCGCTGAGTTCCTGGGGAATGTC 711
DB |||||
661 AATGTGGACGACATTCAGACGCGCAGATGAACCTGACTCTGTGAGTTCTTGGGAAATGC 720
QY 712 AACGAACTGAGTGGGAAGACACACCCCGGTGGCCACCGCCAGAACATGCCGGGAC 771
DB |||||
721 AATGAATCTGAGTGGGAAGATGACACCCAGTGGCCACCGCCAAAACATGCTGTGTGAC 780
QY 772 AGCGCGGATCTATTGGGGA-CGGCAGCAGCGAGACGCGCGCGCCCAACGGGCGCT 830
DB |||||
781 AGTGGGACCTGTTTGGGACCGGCTCTCGGAGACGCGCATGTGGCCACGCTGCTGT 840
QY 831 GTGGCGGACAGTGATCATCGGGAGCAAGACACCGTATAGACTGTCATGATCCGGCC 890
DB |||||
841 GTGGCGAACTGTATCATAGGGGAGCAAGAGCATGCTACGACCTGCTATGATCGGCC 900
QY 891 TTACATGAAGTGTCAACCCAGGAGGTACTACGGGAGGCTCAACGCCATCATGCT 950
DB |||||
901 CTACATGAAGTGTCAACCCATGGAGGATACACGGGAGGCTCAACGCCATCATGCT 960

QY 951 CTTCCAGCGCTGCTTCTTCCAGACAGCAGCGCTCCCGCAGCTACCACTACATCATGAGAA 1010
DB |||||
961 GTTTGAGCGCTGCTTCTTCCAGATAGCAGCTCCCGAGACTATCACTACATCATGAGAA 1020
QY 1011 CTTCTTCTGTAGCTCATCAGCAGCTTAGAGCTCTCTGGTGTGAGGACTACATGATCT 1070
DB |||||
1021 TCTCTTCTTATACGTCATCATCAGCTAGAACTGCTCGTGGCTGAGGACTACATGATCT 1080
QY 1071 GTACCTGAACGTCAGCGCCCGCGGAGGATGCTGGAATCGGCTGCTCAAGAGTG 1130
DB |||||
1081 GTATCTGAACGCGCCGACCGCCCGGAGGAGGATGCTGGCATTTGGTGGCTGAAGAGTG 1140
QY 1131 CTACAGATGATCGACCGGAGGTTTCGGAAAAACCTGAAGTCTTGTATCATCTCCACCC 1190
DB |||||
1141 TTACCACATGATTGACAGGAGACTGAGGAGATCTTAAGTCCCTGATCATGCTCCACCC 1200
QY 1191 CTCGTGGTTTCAATCGGACTGTGTGGCCATCTCTCGCCCTTTTCATCAGCGTCAAGTTCT 1250
DB |||||
1201 CTCCTGGTTTCAATCGCACTGTGTGGCCATCTCTCCGGCCATTCATCAGTGTCAAGTTCT 1260
QY 1251 CAACAAGATCCAGTACGTGCAAGCTTGGAAGCTTGGAAGCTTGGAAGCTTGGAAGCA 1310
DB |||||
1261 CAGTAAATTCAGTACGTGCAAGCTTGGAAGCTTGGAAGCTTGGAAGCTTGGAAGCA 1320
QY 1311 CTTCCAGATCCAGACTGCTGCTGCAATAAGAGAGAAAGACTGGAAGGCCAGGAGGA 1370
DB |||||
1321 CGTGAGCTGCGAGACTGTGTCTGCAATATGAAGAGCAGAGACTCCGAGCCAGAGGA 1380
QY 1371 GAGCGCGAGG---CCCCAGCGGAGTTTGTGTGCTGCCAGGTCTGAAGAGAACCCAGAGT 1427
DB |||||
1381 GAGCAGACGGCCACCGCAGCGGAGTCTCTCTTCCAGGTGAGAGAACCCAGAGAC 1440
QY 1428 GGCACCACTGGAAGACAGCTGCTGCTGTCTCAGAGATCAGGAAACAGCATCTCTG 1487
DB |||||
1441 TGTGAAGAGAGGAGCAGGCGCAGAGGCAACAGAGGACCGAGAACTAGCATCTCTG 1500
QY 1488 AGCGCAGCTGAGCATAAACAAAGGACATGGAAGACATGCAAGACATCCAGATGCCAGAACTCTGT 1547
DB |||||
1501 ATCTACCCAGAACCTTAGACATGGA---GAGATTATTTCCCAATCATGTAAACCTCAT 1557
QY 1548 CAGACGCCCATGCGCCCGCAGATCTCATCTGCTCATCC 1586
DB |||||
1558 CGAAGCCCTGGAGCCCGCTTCATCGAGCTCCACC 1596

RESULT 8

DQ053108

LOCUS

DEFINITION

Pan troglodytes Hc13650 gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

DQ053108

VERSION

DQ053108.1 GI:66899055

KEYWORDS

GSS.

SOURCE

ORGANISM

Pan troglodytes (chimpanzee)

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominoidea; Pan.

1 (bases 1 to 996)

Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,

Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeille, D.,

White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

A Scan for Positively Selected Genes in the Genomes of Humans and

Chimpanzees

(er) PLOS Biol. 3 (6), E170 (2005)

15869325

2 (bases 1 to 996)

Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,

Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeille, D.,

White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

Direct Submission

Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

FEATURES
source Location/Qualifiers
1..996
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>996
/locus_tag="HC13650"

gene
ORIGIN

Query Match 26.8%; Score 783; DB 11; Length 996;
Best Local Similarity 78.9%; Pred. No. 1.2e-109;
Matches 786; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 450 GCCACTCCCAAGAGACGGGGTGGAACTGCTTGGCAGCCCGGTGGAGACACATCTTC 509
Db 1 GCCACTCCCAAGAGACGGGGTGGAACTGCTTGGCAGCCCGGTGGAGACACATCTTC 60

QY 510 TCCTCCCAACACGCTAAATTTCAACGAGCGCATCTTAAGAGGAAGACGCTGGTGCCCC 569
Db 61 TCCTCCCAACACGCTAAATTTCAACGAGCGCATCTTAAGAGGAAGACGCTGGTGCCCC 120

QY 570 AGAGATCAACATTTCTTGATCAGATGAGGGTCTCTGCTCGATGACTTCTTGA 629
Db 121 AGAGATCAACATTTCTTGATCAGATGAGGGTCTCTGCTCGATGACTTCTTGA 180

QY 630 TACCCCTGATGACTGATTAACCTGGATGATCGAGACCCCGGTGAGACGACATC 689
Db 181 TACCCCTGATGACTGATTAACCTGGATGATCGAGACCCCGGTGAGACGACATC 240

QY 690 GCTGGAGTTCTTGGGAAATGGCAACGAACTGGAGTGGGAGAGACACACCCCGTGGCCAC 749
Db 241 GCTGGAGTTCTTGGGAAATGGCAACGAACTGGAGTGGGAGAGAGAGAGACACCGTAT 300

QY 750 CGCCAGAACATCCCGGGAGACAGCGGATCTATTGGGAGCGGACGAGAGACGG 809
Db 301 CGCCAGAACATCCCGGGAGACAGCGGATCTATTGGGAGCGGACGAGAGACGG 360

QY 810 CAGCGCCCAACAGCGGCTGTGGCGGACAGTATCATCGGGAGCAAGACACCGTAT 869
Db 361 CAGTGGCCCAACAGCGGCTGTGGCGGACAGTATCATGGGAGCAAGAGACACCGTAT 420

QY 870 AGACCTGCATGATCCGGCTTTACATGAAGTGGTCAACCCAGGAGGTACTACGGCGA 929
Db 421 AGACCTGCATGATCCGGCTTTACATGAAGTGGTCAACCCAGGAGGTACTACGGCGA 480

QY 930 AGGCTCAACGCCATCATGCTTTCGAGCTGCTTCTTCCAGACAGAGCTTCCCGA 989
Db 481 AGGCTCAACGCCATCATGCTTTCGAGCTGCTTCTTCCAGACAGAGCTTCCCGA 540

QY 990 CTACCACTATCATATGAGAACCTCTTCTGTACGTATCATCAGCAGCTTAGCTCTGT 1049
Db 541 CTACCACTATCATGAGAACCTCTTCTGTACGTATCATCAGCAGCTTAGCTCTGT 600

QY 1050 GGCTGAGGACTACATGATCGTGTACCTGAAACGGTGCACGCCCGCGGAGGATGCTGG 1109
Db 601 GGCTGAGGACTACATGATCGTGTACCTGAAACGGTGCACGCCCGCGGAGGATGCTGG 660

QY 1110 AATCGGCTGGCTGAAGAAGTGTACAGATGATCGACCGGAGGTTCCGGAAGAACTGAA 1169
Db 661 AATCGGCTGGCTGAAGAAGTGTACAGATGATCGACCGGAGGTTCCGGAAGAACTGAA 720

QY 1170 GTCTTTGATCATGCTCCACCCCTCGTGGTTTATTCGAGCTGTGCTGCTGCTGCT 1229
Db 721 NNN 780

QY 1230 TTTTCATCAGCGTCAAGTTTCAACAAGATCCAGTACGTGACAGCTTGAAGACCTGGA 1289
Db 781 NNN 840

QY 1290 GCAACTCATCCCTATGGAACACGCTCCAGATCCAGAGCTGGTCTCTGCAATACGAAGAGGA 1349

Db 841 NNN 1409
QY 1350 AAGACTGAAGCGCAGGAGCGCGAGCGCCCGGAGTTTGTCTCCCGAGTC 1409
Db 901 AAGACTGAAGCGCAGGAGCGCGAGCGCCCGGAGTTTGTCTCCCGAGTC 960
QY 1410 TGAAGAGAGCCAGAGGTGGCAACAGTGGAAAAACAG 1445
Db 961 TGAAGAGAGCCAGAGGTGGCAACAGTGGAAAAACAG 996

RESULT 9
CN420201 748 bp mRNA linear EST 16-MAY-2004
328775518 GRN_ES Homo sapiens CDNA 5', mRNA sequence.
CN420201
CN420201.1 GI:47407795
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 748)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 748 Std Error: 0.00.

FEATURES
source Location/Qualifiers
1..748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_lib="GRN_ES"
/note="oligo dt primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Query Match 25.6%; Score 748; DB 7; Length 748;
Best Local Similarity 100.0%; Pred. No. 2.9e-104;
Matches 748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GCTACTCGGAGCG 96
Db 1 GCTACTCGGAGCG 60

QY 97 CGCGCTGGGAGCG 156
Db 61 CGCGCTGGGAGCG 120

QY 157 GGGCGCTGGCGACAGTGGCTGGGTGACCTTCTCGATGCAAGATTCGCGCCCTCGGAGC 216
Db 121 GGGCGCTGGCGACAGTGGCTGGGTGACCTTCTCGATGCAAGATTCGCGCCCTCGGAGC 180

QY 217 ATCTCTTCTCTAGGCTCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 276
Db 181 ATCTCTTCTCTAGGCTCTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240

```
QY 277 GGCAGGCTCGCTCTTTGTTGGCCAGTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGTTC 336
Db 241 GGCAGGCTCGCTCTTTGTTGGCCAGTAAGGAGGAGGCTGTCTCAGCTGCAGAGGGTTC 300
QY 337 ATCCCTGCTTCAAGCCAGTGCCTCTTCCAGCTCCCATGGGACACACCGAAGCAGCTC 396
Db 301 ATCCCTGCTTCAAGCCAGTGCCTCTTCCAGCTCCCATGGGACACACCGAAGCAGCTC 360
QY 397 CGGATGGAAACCTGTGACGTGAAGGAGGAATGGCAGGACGAAGATCTTCCAGGCCACTC 456
Db 361 CGGATGGAAACCTGTGACGTGAAGGAGGAATGGCAGGACGAAGATCTTCCAGGCCACTC 420
QY 457 CCAGAAGAGACGGGGTGGAACTCTTGGCAGCCCGGCTGGAGACACATCTCTCTCC 516
Db 421 CCAGAAGAGACGGGGTGGAACTCTTGGCAGCCCGGCTGGAGACACATCTCTCTCC 480
QY 517 AACACGCTAAATTTCAACGAGCGCATCGTAAGAGGAAGACGCTGTGGCCCGCAGAGATC 576
Db 481 AACACGCTAAATTTCAACGAGCGCATCGTAAGAGGAAGACGCTGTGGCCCGCAGAGATC 540
QY 577 AACATTTCTGTGATCAGAGTGAAGGGTCCCTGCTGTCCGATGACTTCTTGGATACCCCT 636
Db 541 AACATTTCTGTGATCAGAGTGAAGGGTCCCTGCTGTCCGATGACTTCTTGGATACCCCT 600
QY 637 GATGACCTGGATTAACGTGTGATGATCGAGACCCCGATGAGACCGACTCGCTGGAG 696
Db 601 GATGACCTGGATTAACGTGTGATGATCGAGACCCCGATGAGACCGACTCGCTGGAG 660
QY 697 TTCTCTGGGGAATGGCAACCACTGGAGTGGGAAGACGACACCCCGTGGCCACCGCAG 756
Db 661 TTCTCTGGGGAATGGCAACCACTGGAGTGGGAAGACGACACCCCGTGGCCACCGCAG 720
QY 757 AACATGCCCGGGACACGCGGATCTAT 784
Db 721 AACATGCCCGGGACACGCGGATCTAT 748
```

RESULT 10

BG912726
LOCUS 602807930P1 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4940286
DEFINITION 5', mRNA sequence.

ACCESSION BG912726

VERSION BG912726.1 GI:142933202

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 865)

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10878 row: k column: 07

High quality sequence stop: 821.

Location/Qualifiers

1..865

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4940286"

/tissue_type="anaplastic oligodendroglioma with lp/19q

loss"

FEATURES

source

RESULT 11

CV024962

LOCUS

DEFINITION 2512 Full Length cDNA from the Mammalian Gene Collection Homo

CV024962

LOCUS

DEFINITION 2512 Full Length cDNA from the Mammalian Gene Collection Homo

CV024962

LOCUS

DEFINITION 2512 Full Length cDNA from the Mammalian Gene Collection Homo

/lab host="DH10B (T1 phage-resistant)"

/clone lib="NCI CGAP_Brn67"

/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.

Average insert size 2.3 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

	Query Match	25.0%	Score 730;	DB 2;	Length 865;
	Best Local Similarity	95.7%	Pred. No. 1.5e-101;		
	Matches 804;	Conservative 0;	Mismatches 30;	Indels 6;	Gaps 5;
QY	839	CAGTGATCATCGGGAGCAAGAGCACCGGTATAGACCTGCACATGATCCGGCTTACATCA	898		
Db	1	CAGTGATCATCGGGAGCAAGAGCACCGGTATAGACCTGCACATGATCCGGCTTACATCA	60		
QY	899	AAAGTGGTCAACCCAGCGAGGGTACTACGGCGAAGGGCTCAACGCCATCATCGTCTTCG	958		
Db	61	AAAGTGGTCAACCCAGCGAGGGTACTACGGCGAAGGGCTCAACGCCATCATCGTCTTCG	120		
QY	959	CCTGCTTCTTCCAGACAGCAGCCTCCCGACTACCACTACCATCATGGAAGACCTCTTC	1018		
Db	121	CCTGCTTCTTCCAGACAGCAGCCTCCCGACTACCACTACCATCATGGAAGACCTCTTC	180		
QY	1019	TGTACGTCAATCAGCAGCTTATAGAGCTCTGGTGGCTGAGGACTACATGATCGTGTAC	1078		
Db	181	TGTACGTCAATCAGCAGCTTATAGAGCTCTGGTGGCTGAGGACTACATGATCGTGTAC	240		
QY	1079	ACGCTGCCACGCGCCCGCGGAGGATGCTTGGAAATCGGTGGCTGAAGAGTCTACACGA	1138		
Db	241	ACGCTGCCACGCGCCCGCGGAGGATGCTTGGAAATCGGTGGCTGAAGAGTCTACACGA	300		
QY	1139	TGATCGACCGGAGGTTGGGAAACCTGAAAGTCTTTCATCATCGTCCACCCCTCGTGT	1198		
Db	301	TGATCGACCGGAGGTTGGGAAACCTGAAAGTCTTTCATCATCGTCCACCCCTCGTGT	360		
QY	1199	TCATTCGGAGCTGTCTGGCCCATCTCTCGCCCTTTCATCAGGCTCAAGTTTCAACAAGA	1258		
Db	361	TCATTCGGAGCTGTCTGGCCCATCTCTCGCCCTTTCATCAGGCTCAAGTTTCAACAAGA	420		
QY	1259	TCCAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCCCTATGGAACACGTC	1318		
Db	421	TCCAGTACGTGCACAGCTTGGAAAGACCTGGAGCAACTCATCCCTATGGAACACGTC	480		
QY	1319	TCCAGTACGTGCCTGCAATACGAGAGGAAGACTGAAGCCAGGAGGAGAGCGCA	1378		
Db	481	TCCAGTACGTGCCTGCAATACGAGAGGAAGACTGAAGCCAGGAGGAGAGCGCA	540		
QY	1379	GGCCCCAGCGGAGTTTGTGTGCTGCCAGGTCTGAAGAGAAGCCAGAGTGGCACCAGT	1438		
Db	541	GGCCCCAGCGGAGTTTGTGTGCTGCCAGGTCTGAAGAGAAGCCAGAGTGGCACCAGT	599		
QY	1439	AAAAACAGTCTGTCTGTGTCTCAGAAGATCAGGAAACAAAGCATGTCTGAGCGCAGTGA	1498		
Db	600	AAAAACAGTCTGTCTGTGTCTCAGAAGATCAGGAAACAAAGCATGTCTGAGCGCAGTGA	659		
QY	1499	GCATAACAAAGACATGGAAGAGATTCAGAA--TGCCAGAAAACCTCTGTGACAGCGCC	1556		
Db	660	GCATAACAAAGACATGGAAGAGATTCAGAA--TGCCAGAAAACCTCTGTGACAGCGCC	719		
QY	1557	ACTGGCCCCAGATCTCATCTCCCTGCTCATCTGAGTCCCAATCTTCCAGGGTGGCAG	1616		
Db	720	ACTGG--ACCAGATCTCATCTCCCTGCTCAT--CTGAGTCCCATCTTCCAAAGGGTGG	776		
QY	1617	CTCCGTTCATCTCTGAAACCCAGCATCTTTTCAGCTGCTTTGAAAAATTTTATTTT	1676		
Db	777	CTCCGTTCATCTCTGAAACCCAGCATCTTTTCAGCTGCTTTTACAGTGGTTGAACACT	836		

RESULT 11

CV024962

LOCUS

DEFINITION 2512 Full Length cDNA from the Mammalian Gene Collection Homo

CV024962

LOCUS

DEFINITION 2512 Full Length cDNA from the Mammalian Gene Collection Homo

CV024962

LOCUS

DEFINITION 2512 Full Length cDNA from the Mammalian Gene Collection Homo

sapiens cDNA 5' similar to BC026217, mRNA sequence.
CV024962
VERSION CV024962.1 GI:51482851
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 699)
AUTHORS Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S.,
Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O.,
Clingingsmith, T.R., Hartley, J.L., Espósito, D., Cheo, D., Moore, T.,
Simmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C.,
Vandenhaute, J., Cusick, M.E., Albaladejo, J.S., Hill, D.E. and Vidal, M.
Human ORFome Version 1.1: a Platform for Reverse Proteomics
Genome Res. (2004) In press
Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
ORF Sequence (ag (ost) of Gateway Entry construct. Each cloned ORF
results from a PCR reaction using an MGC full-length cDNA as
template DNA and ORF specific primers
PCR Primers
FORWARD: ATGGGACACCGAGCCAC
BACKWARD: CAGGACATCGTTGTTCTGAT
Insert length: 699 Std Error: 65.00
Plate: 11058 row: 09 column: E
Seq primer: ACTGGCGCTGTTTACAACTCGTGACTGGGAAC
High quality sequence start: 98
High quality sequence stop: 698
POLYA=No. Location/Qualifiers
1..699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="mixed"
/clone_lib="Full Length cDNA from the Mammalian Gene
Collection"
/notes="Vector: mixed; The ORFs were PCR amplified from the
MGC (Mammalian Gene Collection) as of April 2004 and
cloned by recombinational Gateway cloning into pDONR223
Donor vector. Reference : MGC (Mammalian Gene Collection)
Program Team, Generation and Initial Analysis of more than
15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
2002, 99(26), 16899-16903"

Query Match 23.9%; Score 696.4; DB 7; Length 699;
Best Local Similarity 99.7%; Pred. No. 2.1e-96;
Matches 697; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 373 ATGGGACACCGAGCCAGCTCCCGATGGAAACCTGGAGCTGAAGCAGGAATGGCAG 432
DB 1 ATGGGACACCGAGCCAGCTCCCGATGGAAACCTGGAGCTGAAGCAGGAATGGCAG 60
QY 433 GACGAAGATCTTCCAGGCCACTCCAGAGAGACGGGGGTGGAACTGCTGGCAGCCCG 492
DB 61 GACGAAGATCTTCCAGGCCACTCCAGAGAGACGGGGGTGGAACTGCTGGCAGCCCG 120
QY 493 GTGGAAGACATCTCTCTCCCAACACGCTAAATTTCAAACGAGCGCATCGTAGAGG 552
DB 121 GTGGAAGACATCTCTCTCTCCCAACACGCTAAATTTCAAACGAGCGCATCGTAGAGG 180
QY 553 AAGACGCTGTGGCCCGCCAGAGATCAACATTTCTCTGATCAGATGAGGGGTCCTCGTG 612
DB 181 AAGACGCTGTGGCCCGCCAGAGATCAACATTTCTCTGATCAGATGAGGGGTCCTCGTG 240

QY 613 TCCGATGACTTCTTGGATACCCCTGATGACCTGTGATATTAACTGGATGACATCGAGACC 672
DB 241 TCCGATGACTTCTTGGATACCCCTGATGACCTGTGATATTAACTGGATGACATCGAGACC 300
QY 673 CCCGATGAGACCGACTCGCTGGAGTTCTTGGGAATGGCAACGAACTGGAGTGGGAAGAC 732
DB 301 CCCGATGAGACCGACTCGCTGGAGTTCTTGGGAATGGCAACGAACTGGAGTGGGAAGAC 360
QY 733 GACACCCCGTGGCCACCGCAAGAACATGCCCGGGGACAGCGCGGATCTATTTCGGGAC 792
DB 361 GACACCCCGTGGCCACCGCAAGAACATGCCCGGGGACAGCGCGGATCTATTTCGGGAC 420
QY 793 GGCACGACGGAGGACGGCAGCGCCCAACGGGCGCCTGTGTGGCGGACAGTGTATTCGGG 852
DB 421 GGCACGACGGAGGACGGCAGCGCCCAACGGGCGCCTGTGTGGCGGACAGTGTATTCGGG 480
QY 853 GAGCAAGAGCACCCTATAGACCTGCACATGATCCGGCCTTATCATGAAGTGTGTACCCAC 912
DB 481 GAGCAAGAGCACCCTATAGACCTGCACATGATCCGGCCTTATCATGAAGTGTGTACCCAC 540
QY 913 GGAGGGTACTACGGCGAAGGCTCAACGCCATCATCTTTCGACGCTTCTTCCATTCCTCCA 972
DB 541 GGAGGGTACTACGGCGAAGGCTCAACGCCATCATCTTTCGACGCTTCTTCCATTCCTCCA 600
QY 973 GACAGCAGCTTCCCGACTTACCATCATCATGAGAACTCTTCTGTACGTCATCAGC 1032
DB 601 GACAGCAGCTTCCCGACTTACCATCATCATGAGAACTCTTCTGTACGTCATCAGC 660
QY 1033 AGCTTAGAGCTCTCGTGGCTGAGGACTACATGATCGTG 1071
DB 661 AGCTTAGAGCTCTCGTGGCTGAGGACTACATGATCGTG 699

RESULT 12
LOCUS DR002011
DEFINITION TC120419 Human fetal brain, large insert, pCMV expression library
Homo sapiens cDNA clone TC120419 5' similar to Homo sapiens
K18A1872 protein (K18A1872), mRNA sequence.
ACCESSION DR002011
VERSION DR002011
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 714)
AUTHORS Birckett, C., Cho, J., Gau, Y., Hamer, R., Kelly, S., Kovacs, K., Liu, L.,
Liu, X., Porter, J., Sachs, A., Shu, Y., Sun, Z., Wong, J., Wu, M.,
Zhang, X., Jay, G. and He, W.
High-throughput cloning of full-length human cDNAs directly from
cDNA libraries optimized for large and rare transcripts
Unpublished (2005)
Contact: Kovacs, KF
High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cDNA@origene.com
This EST submission is part of an on-going human full-length
cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pCMV6 5prime forward vector primer, Origene
Technologies Inc.
Location/Qualifiers
1..714
FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC120419"
/issue_type="Fetal Brain"
/clone_lib="Human fetal brain, large insert, PCMV
expression library"
/note="Organ: Fetal Brain; Vector: pCMV6-XL4; Site 1:
EcoRI; Site 2: XhoI/SalI compatible end ligatio; Oligo-dT
primed reverse transcription optimized for large and GC
rich mRNA transcripts. cDNA size selection, optimized
ligation for large inserts into mammalian expression
vector, random clones selected for end sequence
verification of full-length genes"

ORIGIN

Query Match 23.8%; Score 693.4; DB 8; Length 714;
Best Local Similarity 99.9%; Pred. No. 6.1e-96;
Matches 694; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 49 GCAGGCGCGCAGGCGGCGGCGATGAGAGGGGGCGGCGAGCGCGCGCGCTGGGGAG 108
Db 20 GCAGGCGCGCAGGCGGCGGCGATGAGAGGGGGCGGCGAGCGCGCGCGCTGGGGAG 79
QY 109 CCCACGGCTAACCTGTGACCCACCCACCCCTGTGCACAAAGAGCTGGGGCGCTGGCCA 168
Db 80 CCCACGGCTAACCTGTGACCCACCCACCCCTGTGCACAAAGAGCTGGGGCGCTGGCCA 139
QY 169 CGTCCGCTGGGTGACCTTCTCGGATGAGATCGGCGGCGGCGGCGGCGGCGGCTTCTCTC 228
Db 140 CGTCCGCTGGGTGACCTTCTCGGATGAGATCGGCGGCGGCGGCGGCGGCGGCTTCTCTC 199
QY 229 CTAGGCTCTGAAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 288
Db 200 CTAGGCTCTGAAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 259
QY 289 CTTTGTTCGCCAGTGAAGAGAGAGCTCTCTCAGCTGAGAGGGGTATCCCTCTCTTCA 348
Db 260 CTTTGTTCGCCAGTGAAGAGAGAGCTCTCTCAGCTGAGAGGGGTATCCCTCTCTTCA 319
QY 349 AGCCAGTGCTCTTCCAGCTCCCATGGGACCGGAGCCAGCGCTCCGATGGAAGAC 408
Db 320 AGCCAGTGCTCTTCCAGCTCCCATGGGACCGGAGCCAGCGCTCCGATGGAAGAC 379
QY 409 GTGAGCGTGAAGAGAGAGATGGCAGACGAAGATCTTCCAGCGCACTCCCAAGAGAGAG 468
Db 380 GTGAGCGTGAAGAGAGAGATGGCAGACGAAGATCTTCCAGCGCACTCCCAAGAGAGAG 439
QY 469 GGGGTGGAATCTGTTGGCAGCCCGGTGGAGACACATCTCTCTCCCAACACGCTAAAT 528
Db 440 GGGGTGGAATCTGTTGGCAGCCCGGTGGAGACACATCTCTCTCCCAACACGCTAAAT 499
QY 529 TTCAACGAGCGCATGTAAGAGAGAGAGCTGGTGGGCGGCGGCGGCGGCGGCGGCGG 588
Db 500 TTCAACGAGCGCATGTAAGAGAGAGAGCTGGTGGGCGGCGGCGGCGGCGGCGGCGG 559
QY 589 GATCAGAGTGAAGGGTCCCTGTCTGTCGATGACTTCTTGGATACCCCTGATGACCTGGAT 648
Db 560 GATCAGAGTGAAGGGTCCCTGTCTGTCGATGACTTCTTGGATACCCCTGATGACCTGGAT 619
QY 649 ATTAACGTGGATGATACGAGACCCCGATGAGACCGACTCGCTGGAGTTCCTGGGGGAT 708
Db 620 ATTAACGTGGATGATACGAGACCCCGATGAGACCGACTCGCTGGAGTTCCTGGGGGAT 679
QY 709 GCGAAGCAACTGGAGTGGGAAGACGACACCCCGT 743
Db 680 GCGAAGCAACTGGAGTGGGAAGACGACACCCCGT 714

RESULT 13

BG913524
LOCUS 876 bp mRNA linear EST 05-JUN-2001
DEFINITION 602811270F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4943369

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

5', mRNA sequence.
BG913524
BG913524.1 GI:14294000
EST.
Homo sapiens (human)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10886 row: k column: 18
High quality sequence stop: 693.
Location/Qualifiers
1..876

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4943369"
/issue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Brn67"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 22.4%; Score 653; DB 2; Length 876;
Best Local Similarity 94.5%; Pred. No. 8.3e-90;
Matches 774; Conservative 0; Mismatches 35; Indels 10; Gaps 9;
QY 1788 CCGCGTGGCTGGTGGCTCCGCTCTGCTCGAGCCCTGTGGTGGTGGTGGTGGTGGTGGTGG 1847
Db 1 CCGCGTGGCTGGTGGCTCCGCTCTGCTCGAGCCCTGTGGTGGTGGTGGTGGTGGTGGTGG 60
QY 1848 AAGATTCAAGACCCCTTCTTGTGCTTGTCAACCCGCTCCAGGTTGGAGCCACAGACCCAC 1907
Db 61 AAGATTCAAGACCCCTTCTTGTGCTTGTCAACCCGCTCCAGGTTGGAGCCACAGACCCAC 120
QY 1908 CGCCACCCCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1966
Db 121 CGCCACCCCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
QY 1967 GACCTAGAGCTCAACCCCTTATGAGGCGCAGCTCTGCGGGTAGCTCTGAGCTCCGAC 2026
Db 181 GACCTAGAGCTCAACCCCTTATGAGGCGCAGCTCTGCGGGTAGCTCTGAGCTCCGAC 240
QY 2027 CTTATGTCACAAATTTACACCCATGGTTTTCATTTGACCGCCCTTCTCGCTCATAA 2086
Db 241 CTTATGTCACAAATTTACACCCATGGTTTTCATTTGACCGCCCTTCTCGCTCATAA 300
QY 2087 TGACACCCAGCTCTCTTTGAGAGGATCAGAGCCCATTTGCAAGAGAGCCGCTGCCAACC 2146
Db 301 TGACACCCAGCTCTCTTTGAGAGGATCAGAGCCCATTTGCAAGAGAGCCGCTGCCAACC 360
QY 2147 ATCCTTGTCTCCGATTGCAAAATGACACCCAGTAATCTAGAACATTTCTCAAGCCCTT 2206
Db 361 ATCCTTGTCTCCGATTGCAAAATGACACCCAGTAATCTAGAACATTTCTCAAGCCCTT 420
QY 2207 TAACTCAGATGTCAGGCCACCGGGCAACCCCGTCAATACCTCCACCAAGATGAGAT 2265

Email: rbrandenberger@eron.com
Insert Length: 639 Std Error: 0.00.
Location/Qualifiers
1. .639

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN PRENEU"
/notes="oligo dt primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."

ORIGIN

Query Match 21.9%; Score 639; DB 7; Length 639;
Best Local Similarity 100.0%; Pred. No. 1.2e-87;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 GGGGCGCGACGGCGGGCGGATGAGAGGGGGCGCGAGCGCGCGCGCTGGGGAGCCC 111
Db 1 GGGGCGCGACGGCGGGCGGATGAGAGGGGGCGCGAGCGCGCGCGCTGGGGAGCCC 60

QY 112 ACCGCTAACCTCGACCCCAACCCACCCCTGCAACAAAGAGCTGGGGGGCGCTGGCCACGT 171
Db 61 ACCGCTAACCTCGACCCCAACCCACCCCTGCAACAAAGAGCTGGGGGGCGCTGGCCACGT 120

QY 172 CGCCCTGGGTGACCTTCTCGGATGAGATCGCCCTCGGAGCATCTCTTCCTCCTA 231
Db 121 CGCCCTGGGTGACCTTCTCGGATGAGATCGCCCTCGGAGCATCTCTTCCTCCTA 180

QY 232 GGCTCTGAAGCGCGGGGAGCGTGAGCGATGCCCGAGCTGCACCGGGCGAGGGCTCGCCTT 291
Db 181 GGCTCTGAAGCGCGGGGAGCGTGAGCGATGCCCGAGCTGCACCGGGCGAGGGCTCGCCTT 240

QY 292 TGTTCGCAAGAGGAGGAGAGCGTGTCTCAGCTGAGAGGGGTATCCCTGCTTCAAGC 351
Db 241 TGTTCGCAAGAGGAGGAGAGCGTGTCTCAGCTGAGAGGGGTATCCCTGCTTCAAGC 300

QY 352 CAGTGCTCTTCCAGCTCCCATGGGAGCACCGAGCCACGCTCCGGATGGAAACGTG 411
Db 301 CAGTGCTCTTCCAGCTCCCATGGGAGCACCGAGCCACGCTCCGGATGGAAACGTG 360

QY 412 GAGCTGAAGGAGGAATGGCAGGACGAAGATCTTCCAGGCCACTCCAGAGAGAGCGGG 471
Db 361 GAGCTGAAGGAGGAATGGCAGGACGAAGATCTTCCAGGCCACTCCAGAGAGAGCGGG 420

QY 472 GTGGAACTGTTGGCAGCCCGGTGGGAAGACATCTCTCTCCCAACGCTAAATTTTC 531
Db 421 GTGGAACTGTTGGCAGCCCGGTGGGAAGACATCTCTCTCCCAACGCTAAATTTTC 480

QY 532 AACGGAGCGCATCGTAAGGAGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTTGGAT 591
Db 481 AACGGAGCGCATCGTAAGGAGAGAGCGTGGTGGCCCGCAGAGATCAACATTTCTTGGAT 540

QY 592 CAGAGTGAGGGGTCTCTGCTGCTCCGATGACTTCTTGGATACCCCTGATGACCTGGATAT 651
Db 541 CAGAGTGAGGGGTCTCTGCTGCTCCGATGACTTCTTGGATACCCCTGATGACCTGGATAT 600

QY 652 AACGTGGATGACATCGAGACCCCGGATGAGACCGACTCG 690
Db 601 AACGTGGATGACATCGAGACCCCGGATGAGACCGACTCG 639

Search completed: March 5, 2006, 06:52:44
Job time : 11158 secs

THIS PAGE BLANK (USPTO)